

More extraterrestrial hoopla

The grand new space policy of the United States is philosophically coherent but intellectually bankrupt, promising little of practical benefit to those left on the Earth.

THE tragic Challenger accident two years ago seems not, after all, to have imbued Washington's policy-makers with the sobriety that might have been expected. The loss of a whole shuttle crew should have been an occasion for reflection, for sober consideration of what purposes are served by the exploration of space and for a deliberate decision about the framework in which future activity will be contained. In the event, the outcome of a year's brooding is that hardware takes precedence over software. There is to be a base on the Moon and, next century, an excursion to Mars, but meanwhile new lifting equipment will be developed to service these expeditions. At the same time, the government plans to privatize the business of launching Earth satellites, using public funds to enable industrial companies to establish themselves in these capital-intensive ventures. The plan also legitimately centres around the principle that the United States will not neglect its national security in deciding what to do next, and next after that; the question inadequately answered is how national security is secured by the plan now on the table. Does it equate with "leadership" in the sense in which that is most often understood in connection with space developments, the knack of being first?

The privatization proposals nevertheless have the merit of constitutional consistency. The US government differs from most others with ambitions (however modest) in this field by the extent to which practical development and even operation of space machines is contracted out. The admirable, even enviable, consequence is that there are skilled private organizations that can be quickly regrouped to meet the needs of new customers for novel machines and services. The theory, conforming with the equally admirable doctrine that government should be as unobtrusive as possible, is that there is enough skill already in being for providers of launching services to assemble the appropriate skills and to sell them on to potential customers. That, in other fields, is usually true. But will the risk-bearing free-wheeling economy of the United States generate the market necessary to sustain privatized space launching? Not, it appears, just now, which is why the administration plans to stand proxy for customers undeclared to the tune of 70 per cent. So much is understandable. There are plenty of organizations wishing to launch communications satellites, but which are probably unwilling also to have to pay for the launching costs of a complicated private company. A decade from now, it may be different. Meanwhile, what the US administration is attempting is to hurry on the future. The price of that impatience will be to skew the pattern of a future industry, and at great cost. Patience would have won greater benefits.

Lunar adventure

The plan to go back to the Moon, and from there to Mars, is similarly a plan to force the natural pace. It is an adventure in high-technology exploration strictly comparable with the President Kennedy's Apollo programme of the 1960s and, for that matter, with the dash to the South Pole by Scott and Amundsen earlier in the century. As on those occasions, the hard work will be the preparation, but the risks of failure will be great enough to give the expedition an heroic air. And the benefits? More

than a quarter of a century will have passed between the last and the next landing by a person on the Moon, comparable with but even smaller than the interval between the two first journeys to the South Pole and the next visit. Those heroic journeys served chiefly to show that the risks of repetition were not worth the trouble, chiefly because the means of transport were too costly and insecure — and there was so little to do on arrival.

The plan to go to Mars is a plan to make a pageant for a future generation. It is partly inspired by the disappointment of Challenger and partly by the advertised interest in Mars of the Soviet Union. Nearer the time, no doubt, the community of planetary scientists will be asked for its views, and there will be hastily arranged competitions for people wishing to design programmes of observation that might be made above, below and on the surface of that inhospitable place. But neither those who put in bids nor those who pay for them (the taxpayers of that decade) should be disappointed if the programme of martian exploration is quickly put on ice after the first few successes. As at the cancellation of the Apollo programme, the word will go out that it would be more prudent to have a safer means of transport (whence, of course, the shuttle). None of this means that the journeys will not be fun, but simply that a huge investment of resources stretching over decades is justified largely by the undoubtedly absorbing interest of the long preparation, the sense of making progress that will come from that and the no doubt even more absorbing television programme it will be possible to mount from Mars. Naturally, it will be a disappointment for the United States if the Soviet Union does all this first... but there is always even less hospitable Venus.

By-products

The Apollo people were fond of saying that their project would cost a lot of money but that none of it would leave the ground. The same argument will arise on this occasion, and there is much in the claim that the tangible end-product of the mission to Mars will be the skill of the people who will have built the transport system, and which can then be turned in other directions. But as with all development projects whose chief benefits turn out to be their by-products, it would make more sense to mount projects aimed at the by-products themselves.

If, as on this occasion, it is felt necessary to bring about a substantial improvement and broadening of the capacity of the United States to send objects into space (a legitimate component of national security), that could more efficiently (and cheaply) be done by concentrating people's efforts, in the next decade or so, on the jobs that cry out for doing. The careful collection of data from the Solar System nearby, the further development of telecommunications and remote sensing satellites whose work must bring immediate benefits and the development of machines in which reliability takes precedence over spectacle. The temptation is to suppose that such dull work cannot contain the seeds of greatness, let alone leadership. But in this new and risky business, the prizes will go to those who do what they say with unsensational reliability. Columbus, after all, is remembered still not because he was a daring fellow, but because he was a splendid navigator. □



Reagan's new space policy contains little science

- Hopes for a commercial launcher
- Military matters remain a priority

Washington

THE White House last week released details of a new comprehensive space policy signed by President Ronald Reagan at the start of January.

As previously reported, a large part of the policy is aimed at encouraging a commercial space sector to complement the civil and defence space sectors (see *Nature* 331, 380; 1988). It also establishes the long-range goal of expanding human presence beyond Earth orbit, first to a lunar base and ultimately to Mars.

Towards that end, a new technology and development programme, dubbed Project Pathfinder, is due to be started in the 1989 fiscal year beginning next October.

Although the new policy explicitly endorses the construction of a permanently manned space station, it also calls on the National Aeronautics and Space Administration (NASA) to solicit proposals for a different orbiting space facility "suitable for research and commercial manufacturing" to be built, operated and paid for by the private sector. NASA administrator James Fletcher says NASA will seek proposals at once, and expects a contract to be signed by July of this year. Only one company, Space Industries Inc., of Houston, Texas, is known to be ready with a proposal.

Although this Industrial Space Facility (ISF) is intended as a private venture, it will nevertheless depend heavily on government money. Fletcher says the government plans to lease 70 per cent of

it once it is in service. Space Industries reckons this income, worth approximately \$700 million over five years, is necessary to make the project a financial reality.

In addition to ISF, the Reagan administration will also support development of "Spacehab", a commercially owned and operated shuttle module that would also permit human operations in space.

The new policy will take the US government permanently out of the commercial launch business, unless a commercial payload has a specific need for the lift capabilities of the space shuttle. This reiterates a policy first articulated in 1984 putting the Department of Transportation in charge of encouraging a commercial launch industry. Private industry is also being encouraged to take advantage of space aboard the space station once it is launched.

The new policy also makes it quite clear that space activities are needed to "strengthen the security of the United States". Although the United States remains committed to exploration of space for "peaceful purposes", it nonetheless insists that this allows the pursuit of national security goals.

NASA will retain responsibility for space science. But with the financial burdens of building a new shuttle to replace the Challenger, constructing the space station, developing a new heavy lift vehicle, and supporting the new ISF, even the healthy budget increases expected in the 1989 budget will leave precious little for new scientific activities. **Joseph Palca**

ESA and NASA get it together

Paris

Two areas of uncertainty left over from last November's ministerial meeting of the European Space Agency (ESA), in The Hague, regarding Europe's long-term space plan and programmes seem now to have been settled (see *Nature* 330, 302; 1987). The news, like the curate's egg, is good in parts.

The good news is that a compromise has finally been reached with the US National Aeronautics and Space Administration (NASA) regarding joint plans for the \$20,000 million international space station (see *Nature* 331, 469; 1988). ESA, which wants to contribute a 'package' of orbiting laboratories — called Columbus — to interface with a manned US space platform, could not accept that, as major sponsor, the United States should have responsibility for administration of the entire project.

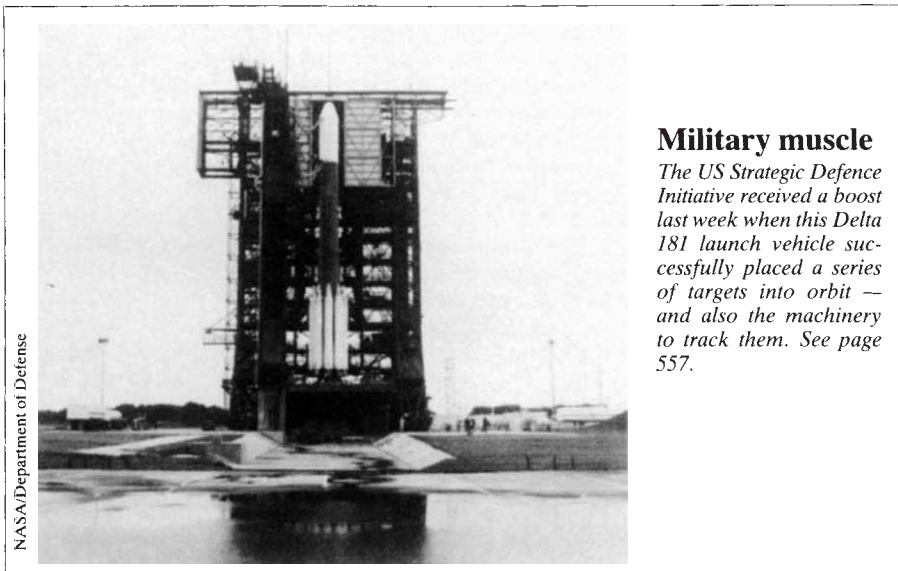
After a week of negotiations in Washington in early February, ESA delegates returned with outlines for a compromise. While the United States will have the final say in decisions relating to the space platform, other countries will retain authority over the elements they contribute. This means that ESA member states will be able to veto use of the Attached Pressurised Module and Man-Tended Free-Flying Laboratory for military experiments. But the United States will still be able to carry out experiments of a military nature within the main space station, so long as these do not involve actual testing of weapons. The ESA charter says its activities must be of a peaceful nature only.

Other bones of contention carried over from The Hague concerned arbitration procedures in the event of a dispute between partners, and patent rights to discoveries arising from microgravity experiments. Although lawyers will now have the difficult task of phrasing the exact terms of a final agreement, US jurisdiction will apply only to the space station.

The bad news, at least for Britain's space industry, is that Mr Kenneth Clarke, Minister of State for Trade and Industry, has confirmed that Britain will not take part in either the Columbus programme or the building of the Ariane 5 heavy-lift rocket launcher.

ESA officials in Paris were reluctant to comment on Britain's non-participation, but hopes were never high that Clarke would change his mind. Member governments now have just over a month to consider the text drawn up in Washington, in time for the next ESA Council meeting on 17 and 18 March.

Peter Coles



Military muscle

The US Strategic Defence Initiative received a boost last week when this Delta 181 launch vehicle successfully placed a series of targets into orbit — and also the machinery to track them. See page 557.

"Profoundly disappointing" UK science budget allocated

London

By choosing to ignore recommendations for a greater investment in the science base, the British government has seriously prejudiced the country's future social and economic development, the Advisory Board for the Research Councils (ABRC) has said in its official reaction to last November's announcement of the science budget (see *Nature* 330, 101; 1987). The board, whose job is to advise the education secretary on his responsibilities for civil science and how to divide the science budget among the five councils, describes as "dismal" the present financial status of British academic science.

The comments are contained in the board's formal advice on the allocation of the science budget for 1988-89 to 1990-91. The budget provides for an increase of £50 million to £699 million for 1988-89, with a planning figure of £729 million for each of the next two years — increases of £65 million and £48 million, respectively, on previous plans.

ABRC chairman Sir David Phillips says in his covering letter to the education secretary, Mr Kenneth Baker, that the board was "profoundly disappointed" that the increases were "insufficient either to avert a reduction in the volume of scientific activity or to allow for the strategic

reshaping of the science base". Comments accompanying the allocation advice itself will leave the government in little doubt of the level of bitter frustration that the scientific community is experiencing. The advice concludes: "The government's revised expenditure plans for the science budget do not provide the means to move our nation's scientific capability towards the twenty-first century. We remain precariously poised at the divide and are beginning to slip in the wrong direction. A great opportunity has been sadly missed."

The board calculates that relative to the government's forecasts of inflation in the economy as a whole, the new science expenditure plans imply increases of 1.3 per cent in 1988-89, a further 1.2 per cent the following year, and a reduction of 3 per cent in 1990-91. Of the extra money, £14 million is earmarked for Antarctic and AIDS research, which if set aside results in a decline in science expenditure of about 0.2 per cent in each of the first two years, with no change in 1990-91. But because of the exceptionally high costs of scientific research, the board expects a 2 per cent reduction in scientific activity for this and between 2 and 3 per cent less in 1990-91, compared with an anticipated

cumulative growth of 8 per cent in the country's gross domestic product. "Basic and strategic scientific research will thereby become an even smaller part of our national effort", the board says.

The ABRC had asked for an extra £103 million, £131 million and £166 million for each of the next three years, principally to restructure academic scientific research and teaching. The board points out that committed expenditures will absorb nearly all the money available, with next year's pay awards estimated at around £30 million.

The board recommends that the Natural Environment Research Council

UK science budget allocation 1988-89 (£million)

Research councils	
Agricultural & Food	58.44
Economic & Social	27.66
Medical	146.68
Natural Environment	90.86
Science & Engineering	366.28
Other bodies	
Royal Society	7.94
Fellowship of Engineering	0.75
Science Policy Studies	0.10
Centre for Exploitation of Science and Technology	0.08
Total	698.79

receives £3 million this year to cover restructuring costs, including targeted redundancies, arising from a reduction in commissioned research from government departments (see below). A further £2 million may be required in 1989-90, the board says, money that "will not buy science". A request from the Medical Research Council for an increased allocation for AIDS-related clinical and pathological studies cannot be met.

In all, the sum available for new initiatives this year is only £7.5 million, the main priority for which is the establishment of interdisciplinary university research centres. The board recommends that the Science and Engineering Research Council (SERC) be allocated £2.8 million in 1988-89, and £5 million and £3 million in the following two years to fund a further two research centres on top of the one in superconductivity which it had already announced (see left).

The board also recommends an allocation of £0.3 million, rising to £4.2 million to the Medical Research Council for its proposal to establish a research centre in toxicology. The Agricultural and Food Research Council should receive an extra £1.5 million in 1989-90 and 1990-91 to spend on research in the universities, the board says.

The board welcomes the government's LINK programme, (see *Nature* 331, 473; 1988) to encourage collaborative research between industry and the academic community. The board recommends that SERC be allocated a further £1.1 million in the coming year and £2 million in later years to support further LINK programmes.

Simon Hadlington

NERC cuts back

London

BRITAIN'S Natural Environment Research Council (NERC) is preparing to implement severe cost-cutting measures in the face of a decline in income from research commissioned by government departments.

If the government accepts recommendations on allocation of the science vote, NERC will receive in 1988-89 £90.6 million, a decline in real value of £2.5 million on this year's grant. The value of commissioned income is expected to fall by £4.1 million, caused by a decline of £4.6 million in government commissions, but offset by an increase of £500,000 in non-government commissions. Other income is expected to fall by £700,000, partly because of withdrawal of the Department of Energy's contribution to the Ocean Drilling Program.

The situation is a particularly bitter pill for NERC to swallow in view of its successful attempts to attract income from non-governmental sources. Such income has trebled in the past three years to around £9 million.

NERC will be looking to shed around 120 jobs from a total workforce of some 2,700, and may have to resort to compulsory redundancies. It will be forced to reduce its support of research in universities by £700,000 as well as scrapping a planned increase of £1.9 million. Building work at the British Geological Survey headquarters in Nottingham will be delayed. No new special topic programmes will start this year, and spending on big computers is deferred. □

Three more centres

London

BRITAIN'S Science and Engineering Research Council (SERC) last week announced a further three interdisciplinary research centres. The universities of Glasgow, Liverpool and Oxford will host the centres for engineering design, surface science and molecular sciences, respectively. The three centres will receive government funding of at least £5 million each over their first six years; the precise sum allocated will depend on the contribution of putative industrial partners in the centres' research programmes. The first research centre, to investigate small devices applications of high-temperature superconductivity, was announced at the end of last year, based at the University of Cambridge.

The concept of university research centres, to provide a focus for the development of research programmes in interdisciplinary topics of strategic importance, has the firm backing of the government. For several months SERC has been considering some 80 bids for centres in seven strategic areas of science. Of the four centres supported, about half the money has come from SERC's own resources, and about half from allocations by the Advisory Board for the Research Councils.

A fifth centre, in synthesis and characterization of semiconductor and novel materials, has been approved in principle to be established at a University of London consortium based at Imperial College. The centre will be set up when enough money becomes available. □

Japanese AIDS scandal over trials and marketing of coagulants

Tokyo

A SCANDAL has erupted in Japan over the government's handling of applications for clinical trials and marketing of heat-treated blood coagulants for haemophiliacs. A former head of the Ministry of Health and Welfare's AIDS (acquired immune deficiency syndrome) research group has admitted delaying approval of clinical trials and marketing so that a Japanese blood-product manufacturer could catch up with foreign competition. As a result, many of Japan's haemophiliacs are thought to have been infected with AIDS.

Japan's first AIDS cases were announced by the ministry in March 1985 but the two patients, both haemophiliacs, had already died, one in 1983. The cases were reported in a medical journal by Dr Takeshi Abe of

Teikyo University who served as head of the ministry's AIDS research group from June 1983 to March 1984. A Tokyo AIDS doctor familiar with the inner workings of the ministry told *Nature* in 1986 that announcement of the arrival of AIDS in Japan was delayed to "make preparations and avoid public panic". But it seems that more devious reasons lay behind the delayed announcement.

Blood coagulants that have been heat-treated to kill viruses such as AIDS were marketed in West Germany in May 1981 and in the United States in March 1983, according to Yukio Yasuda, vice-president of the Tokyo branch of the National Association of Haemophiliacs. In the summer of 1983, the association began a campaign urging the government to introduce the

use of heat-treated blood coagulants in Japan. But it was two years before the products were allowed on the market. In the interim, hundreds of Japanese haemophiliacs, many of them young boys, are thought to have been infected with AIDS.

It is a characteristic feature of AIDS in Japan that most sufferers are haemophiliacs — by last October, 34 of Japan's 59 officially confirmed AIDS patients were haemophiliacs, 20 of whom have died. Furthermore, 930 of 986 known AIDS virus carriers are haemophiliacs and the ministry estimates that about 40 per cent of Japan's 5,000 haemophiliacs, or 2,000 people, are infected with the AIDS virus. The prime infection route is strongly suspected to have been blood plasma imported from the United States for manufacture of blood coagulants.

So why did the ministry take so long to approve marketing of heat-treated coagulants? The ministry issued guidelines for clinical testing in November 1983. But, according to Yasuda, Abe then controlled everything. He designated the companies to carry out clinical tests and the hospitals where tests were to be performed. And the *Mainichi* newspaper recently reported that Abe admitted deliberately delaying approval of clinical testing and marketing so that Japan's Green Cross Corporation (Midori Juji) could catch up with companies such as Baxter Trevanol and Bayer and Hoechst. Midori Juji dominates Japan's blood-product market.

In his defence, Abe argues that he avoided giving late approval to only one company to prevent arousing suspicions of the safety of the company's products among haemophiliacs. But the *Mainichi* has also revealed that Abe received a large donation from Midori Juji to set up a non-profit corporation to promote research on haemophilia. The corporation was founded in July 1986 under Abe's leadership with total funds of ¥110 million (\$850,000). And the corporation headquarters is in a two-room apartment in Tokyo's Itabashi-ku, which, according to the landlord, is empty most of the time.

Baxter Trevanol and Bayer Hoechst also contributed ¥10 million to Abe's non-profit corporation. But, according to the *Mainichi*, one of these foreign companies was at first reluctant to contribute and finally did so only after repeated requests from Abe during the process of trying to get approval for clinical testing. But Abe denies any connection between the funds donated and his control over testing.

The 2,000-member National Association of Haemophiliacs plans to demand compensation from the Ministry of Health and Welfare and the pharmaceutical companies that sold non-heat-treated blood coagulants. But the association has not decided whether to demand compensation from Abe. Yasuda sees as the culprit the whole system. **David Swinbanks**

UK health research threatened by "oppressive" new contracts

London

BRITAIN'S Department of Health and Social Security (DHSS) has incensed recipients of its research contracts by insisting that no results of research financed by the department can be published without permission from the secretary of state. Previously, research contracts required only that any proposed publications be shown to the secretary of state and that "any comments which the secretary of state makes shall be considered by the researcher but the researcher shall nevertheless be free to allow publication to go forward in the original form as he thinks fit".

The new wording is starkly different. Contracts now stipulate that publication is "subject to the prior consent of the secretary of state, which consent shall not be unreasonably upheld". Institutions and individuals receiving DHSS research contracts warn that the new wording is oppressive and unacceptable and that new contracts will not be signed in their present form. The matter has been taken up by professional bodies, including the Royal Society and the Society for Social Medicine.

The DHSS says that the wording of contracts has been changed merely to satisfy legal requirements for publications whose copyright is vested in the Crown, the previous wording having conferred inadequate control over publication, even when legal advice might have been against publication. To that extent, says the DHSS, there is no change in policy and no intention to restrict or suppress publication.

Researchers argue that it is a principle of academic freedom that is at stake, and that they cannot be expected to conduct

research in the knowledge that at the end of the day publication may be censored. More specifically, there are fears that the credibility of DHSS-funded research may be undermined if journal editors or readers do not know if results have been suppressed. Furthermore, the clause will enable the department to delay publication (presently if nothing is heard from the department within 28 days, publication may proceed regardless). Another concern is that under the terms of the new contracts, DHSS-funded workers will find it difficult to find collaborators.

The fact that on its past record there is little to suggest that the department would attempt to interfere with publication is immaterial, say the objectors, who want at the very least to see the contracts contain a more precise form of words that define specific circumstances in which consent to publish can be refused.

Although the question of academic freedom seems unlikely to cut much ice with the government, wider public criticism might. Although most of the health department's externally contracted research (for which it will pay some £16 million this year) involves politically uncontroversial topics, such as hospital design or nursing practice, several programmes (for example, investigations into leukaemia incidence around nuclear power stations, or feasibility studies on contracting services to the private sector) are more sensitive. People are not finding difficulty envisaging circumstances where the government would be grateful for the way that contracts are now worded.

Simon Hadlington

Ireland's new biotechnology venture puts marketing first

Galway

THE government of the Republic of Ireland has cheered at least some academic researchers by setting aside IR£1.5 million to help establish BioResearch Ireland, a new contract research organization designed to place contracts for commercial work with university and research institute laboratories. The government's ambitions were described here last week by Ireland's first Minister for Science and Technology, Dr Sean McCarthy.

By international standards, the Republic of Ireland's universities are small, as are their research budgets, while the present minority government is embarked on a programme to cut public spending that may make things worse. Optimistic biotechnologists now hope they may be exempted.

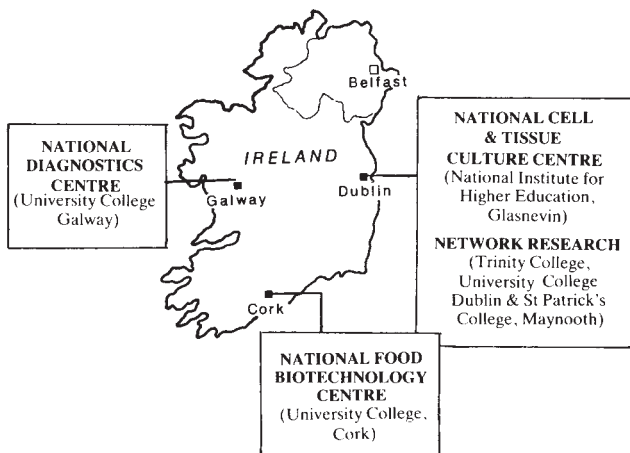
So far, three new research centres have been equipped and staffed at existing laboratories. The National Cell and Tissue Culture Centre, the National Food Biotechnology Centre and the National Diagnostics Centre, already supported by government funds, will now offer, through BioResearch Ireland, contract research and development services to industry. They will also seek industrial collaboration for their own projects.

Barry McSweeney, director of BioResearch, says that part of the IR£1.5 million being provided by the government will be used to set up a system of "network research" to provide researchers not already involved at one of the new centres with equipment and assistance. The two

Dublin universities (Trinity and University colleges) will be drawn in, as will St Patrick's College, Maynooth.

McSweeney says that his organization, which will "commercialize the entire national effort in biotechnology", will be run on strictly commercial lines. Work will be dropped if there is no commercial benefit in prospect, and there will be no direct support for basic research (which nevertheless will benefit indirectly from the availability of equipment), which will remain the responsibility of the Irish Science and Technology Agency, EOLAS.

The agency, formed in January last year by the merger of the previous National Board for Science and Technology and the



Institute for Industrial Research Standards, has a budget of IR£11 million, less by more than IR£1 million than the combined budgets of its predecessors. Prime Minister Charles Haughey's promise during last year's election to cut spending has now materialized in offers of voluntary redundancy to science administrators.

The hope of privatizing a whole sector of science spending has obvious attractions in that context. The next eighteen months should tell whether BioResearch Ireland gains the momentum to accomplish that. So far, only one company has announced a deal with BioResearch. The laboratory instrument maker Beckman is to cooperate on the development of new assay techniques and instrumentation. McSweeney says that other deals are pending but are confidential for the time being, as is his business plan, but that his target for 1988 should be met in the first four or five months.

Ireland may have a trump card to play in attracting industrial partners — the regional development fund of the European Communities, which can help non-Irish companies build factories and otherwise help to establish themselves in Ireland.

Charles Wenz

New biology centre

Two of Britain's wealthiest medical research charities are to set up a major new centre for developmental biology at the University of Cambridge. The Cancer Research Campaign (CRC) and the Wellcome Trust will give the university £4.6 million and £5 million, respectively, to build the centre and run it for five years.

The CRC presently supports a large developmental biology unit in the university's zoology department, under John Gurdon and Ron Laskey. One of the principal reasons for establishing a 'centre of excellence' is to help provide a much-needed focus for talented young scientists who might otherwise be tempted to the United States or into industry. The centre will nominally consist of Wellcome's 'Centre for Developmental Biology' and the CRC's 'Centre for Cell and Molecular Biology', each with its own director.

Wellcome's contribution forms part of its policy to support long-term ventures following an increase in the Trust's income in February 1986 after it sold 21 per cent of its shareholding in the Wellcome Foundation Ltd (now Wellcome plc). The two-year period during which the Trust agreed to sell no more shares has now expired, giving rise to speculation that the release of a further tranche of shares is imminent. S.L.H.

AIDS drug patent

A US patent was granted last week to Burroughs Wellcome for the drug azidothymidine (AZT), marketed as Retrovir, the only drug so far approved for the treatment of AIDS. Patents are also being sought in 40 other countries where the use of AZT has been approved. Burroughs Wellcome, the US arm of Wellcome plc, has no plans to license to any other company the production of AZT, for which it charges about \$8,000 per patient per year. Also last week, the Bristol Myers company received licences to develop two related anti-AIDS drugs, dideoxyadenosine and dideoxyinosine. The National Cancer Institute intends to begin human clinical trials of the new compounds soon. D.L.

Ministers say yes

Two research programmes proposed by the European Community at its ministerial council meeting were approved by the European Parliament on 10 February (see *Nature* 329, 753; 1987). The information technology programme, ESPRIT II will get 3,200 million ECU (about £2,200 million), while BRITE, the technology transfer programme, will get an extra 60 million ECU. With 103 BRITE projects already funded, allowing joint research between universities and industrial partners, the extra grant will enable 66 more projects to start. The European Parliament also agreed the second five-year science programme, with a budget of 165 million ECU. P.C.

Somalis reprieved

Washington

Two Somalis whose plight has been taken up by the US National Academy of Sciences (*Nature* 331, 196; 1988) have had their death sentences commuted to life imprisonment by President Siad Barre. The two, Suleiman Nuh Ali and Abdi Ismail Yonis were among 8 political detainees sentenced to death on 7 February.

The academy last month urged the Somali government to release details about the cases of Ali and Yonis as well as 11 other scientists and engineers all of whom had been imprisoned since 1982. The 13 are accused of political crimes for which little or no evidence has been presented.

Joseph Palca

Presidential election casts shadow on AAAS annual jamboree

Boston

With the presidential election process well underway, the question of how science will fare under a new administration was a dominant theme at the annual meeting of the American Association for the Advancement of Science (AAAS), held here last week.

Although the United States, according to Guyford Stever, former science adviser to President Carter, spends more on basic research than the rest of the non-Socialist world combined, the mood of the meeting was pessimistic.

Major failures of the scientific enterprise were seen as having marred the last few years: the Challenger accident, the tardy and still inadequate response to the epidemic of AIDS (acquired immune deficiency syndrome) and the bitter divisions over the wisdom of pursuing the Strategic Defense Initiative.

No surprise then, that the majority of the speakers appeared to wish themselves back in the 1960s, in the glory days of the Apollo project. Joshua Lederberg, president of the Rockefeller University, called for a revival of the President's Science Advisory Council (PSAC) of that era. Until it was abolished by President Nixon, PSAC, made up of a variety of distinguished leaders from a wide range of scientific disciplines, met periodically to provide advice directly to the president.

Many wished to give the president's science adviser some powerful new teeth. Full cabinet rank was recommended along with the right to sit on the National Security Council and the Economic Policy Council; in other words to take responsibility for science and technology in every sphere of the nation's activity.

The incumbent presidential science adviser, William Graham, made light of all these suggestions. Giving the Science Advisor higher rank would not change anything, he said, for the incumbent is already involved in all relevant decisions. And he could not see the argument for reconstituting PSAC. No substantial difference exists, he said, between PSAC and the present White House Science Council, on which, among others, Edward Teller sits.

William Carey, former executive officer of AAAS, argued that presidential engagement was what mattered most in the quality of science policy. But there may also be a fundamental problem in giving science advice to the president. As Lewis Branscomb, director of the science policy programme at Harvard University, argued, the scientific community wants complete visibility in the advice being given to the president. But the president

requires confidentiality that will enable him to be seen as making his own decisions in a timely fashion, rather than being embarrassed by surprises or public dissent among his advisers. Science Advisors cannot be representatives of science as a constituency, nor can they expect to speak out on matters on which they disagree with the president, however much open debate is a part of the scientific tradition.

There will be plenty of problems for the next Science Advisor to take up. Roland Schmitt, chairman of the National Science Board and until recently vice-president of General Electric, outlined the key issues:

strengthening links between universities and industry without compromising academic freedom; allowing foreign researchers open access without undermining international competitiveness; distributing federal funds more equitably without harming productivity; restructuring industry without harming the research and development base; and finding ways to increase the national laboratories' contribution to economic development.

An issue stressed by William Graham might be added; that of the growing political importance of international cooperation in basic scientific research. Those attending AAAS would have needed no reminder: another whole day of the meeting was devoted to Soviet science and the new possibilities raised for US-USSR cooperative research. **Alun Anderson**

Californians thinking bigger in neuroscience for the 1990s

Los Angeles

THE University of Southern California (USC), whose reputation in the past has rested largely on the quality of its professional schools, has decided to make a pitch for a position among the top US research universities. Its vehicle to excellence is an ambitious interdisciplinary programme in Neural, Informational and Behavioral Sciences (NIBS). The university's stated goal is to be number one in neuroscience by the mid-1990s.

William Wagner, dean for interdisciplinary programmes, says the idea for NIBS was not entirely well received when he

departments in the medical school, and even the law school.

The university has raised \$18 million for a 65,000-square-foot building, scheduled for completion in 1989, that will house the core of the programme. An additional 45,000 square feet in the university's medical school has also been committed to the programme. USC will create 30 new tenured positions for NIBS over the next 5 years, as well as a minimum of 20 junior faculty positions. All members of the faculty will have appointments in one or more existing university departments.

Wagner, a former student of physicist Murray Gell-Mann at the California Institute of Technology, shares Gell-Mann's passion for interdisciplinary programmes, and is looking for people with "renaissance minds", able to apply their perspectives across disciplinary boundaries. Undaunted by the difficulties of attracting faculty to a programme still largely on the drawing board, he has been recruiting aggressively, offering generous salaries and laboratory support.

After 6 years of planning and fund-raising, recruitment is finally gaining momentum, with the arrival of the first 4 new faculty members in the past 6 months. Three more appointments are in the final stages of negotiation, says Wagner.

While neuroscientists praise USC for its efforts, some question its ability to start from scratch and build a neuroscience programme that will rank among the top programmes in the country. Others doubt whether such a broad scope is the best strategy for achieving such excellence. Wagner and those involved in NIBS seem to believe that it is, and much of the programme's fate will ride on their success in persuading others to join them.

Marcia Barinaga



introduced it in 1982. Now that interdisciplinary programmes have come of age, with interest growing in areas such as computational neurobiology, many departments have sought to join the NIBS bandwagon. Up to 100 faculty members may participate, from departments ranging from linguistics to electrical engineering, including all the natural sciences, many

Boost for industrially relevant research in India

New Delhi

THE report of a review committee that recommended revamping of the Council of Scientific and Industrial Research (CSIR) was finally accepted early this month with modifications suggested by the scientific advisory committee to prime minister Rajiv Gandhi. The report was hanging fire for more than a year because of CSIR's objections to the suggested closure of all its 100 extension centres and transfer of four of its laboratories to concerned government departments.

It has now been decided that the laboratories will stay with CSIR and closure of extension centres will be effected on a selective basis. According to CSIR director-general, Dr A.P. Mitra, other recommendations of the committee have been accepted and will be implemented in three months to give CSIR a new look.

Under the new scheme, the CSIR headquarters in Delhi gives up administrative and financial control over its 41 laboratories in favour of greater autonomy. It will confine its role to policy formulation, manpower and technology development. The laboratories will earn one-third of their keep through consultancy and other services to industry and government, the rest coming from the taxpayers.

The proposed structural changes aim at strengthening CSIR's link with industries and user agencies in the government. Henceforth the laboratories will work on research and development projects whose priorities will be identified by a high-level

advisory board headed by an eminent technologist from outside. Besides scientists from CSIR, the board will have members representing industries and user agencies and will be responsible for resource allocations and review of work in progress. Experts from outside CSIR will also be included in half a dozen technical advisory boards which will undertake technology forecasting and formulate mission-oriented projects.

CSIR has also agreed to the committee's recommendation for a fixed six-year term for laboratory direction. Extension will be given in exceptional cases after a review of the director's performance.

An exception made in the case of Dr Pushpa Bhargava, director of the Centre for Cellular and Molecular Biology (CCMB) in Hyderabad, has kindled a controversy. Bhargava, who had exceeded the six-year term and was to attain retirement age on February 28, was first given six weeks' notice to leave, but the order was retracted three weeks later and he was given a two-year extension.

According to CSIR the decision to extend his tenure was taken at the highest level, meaning the prime minister, who reportedly received letters from a section of CCMB scientists and some Nobel laureates, including Francis Crick, urging him to let Bhargava continue. Whether the six other directors who have completed the six-year term will get a similar extension depends on the influence they can bring to bear.

K.S. Jayaraman

Soviet demographers come clean

London

SOVIET planners "fail to see the human being for the worker", according to Professor Dmitri Valentei of the University of Moscow Centre for the Study of Population Problems.

Writing in the daily *Sotsialisticheskaya Industriya*, Valentei complains that responsibility for vital regional developments has too often been left to "dilettantes" who have failed to provide redeployed groups of workers with the necessary social infrastructure, with the result that there has been a drift of population from less favoured areas to the large cities, while smaller towns have declined.

Valentei says that, while the State Commission for Labour of the USSR has now been made responsible for "an effective demographic policy", it simply does not have enough expertise to do the job. Valentei offers his university's services for training people.

The Soviet Union has two urgent demo-

graphic problems — an ageing population and a growing ethnic imbalance. While the Soviet population has increased by a third since 1959, the number of pensioners has increased by 78 per cent.

Life expectancy, which fell in the 1970s, is now increasing again (but Valentei says that it is still "about 10 years for men and 8 years for women less than the best world indices"). There has also been a slight increase of the birth rate.

Valentei also notes the relatively rapid growth of the traditional Islamic islands of Central Asia, Kazakhstan and Azerbaijan, where 18 per cent of the Soviet population appears responsible for 40 per cent of population growth. One dilemma brought to light by *glasnost* is that infant mortality is relatively higher in the Moslem republics (in excess of 30 deaths per 1,000 births), as is the incidence of congenital defects. Steps taken to change this state of affairs could worsen the ethnic imbalance in the short run.

Vera Rich

Embryo research criminal

Munich

IGNORING protests from scientific and medical organisations, the West German cabinet on 10 February recommended criminal penalties for research on human embryos and manipulation of human germ-line cells. Health Minister Rita Süßmuth (Christian Democrat) also recommended that artificial insemination be made illegal for women with infertile husbands and for unmarried couples.

The cabinet broadly followed the counsel of the Benda commission, created by the Justice and Research and Technology Ministries in 1984. But by making embryo research a criminal act, it came into conflict with the Max Planck Society, the Bundesärztekammer (a physicians' professional association) and other scientific organizations.

The cabinet issued a third recommendation forbidding surrogate motherhood — the bearing of a child for a childless couple by a so-called "mother-for-hire". The ban would extend even to the advertising of surrogate mother services. This measure was meant to close a loophole that had allowed the affiliate of a US surrogacy firm to operate until recently in Frankfurt (see *Nature* 329, 577; 1987).

Research organisations agreed in principle to several of the recommendations, for example, to the ban on the creation (cloning) of human beings from totipotent cells or the ban on creating human-animal chimaeras? But Max Planck Society spokesman Peter Gutjahr-Löser protested the "portrayal of research as bordering on criminal activity", saying that such laws "might discourage young researchers" from entering fields such as molecular biology.

The president of the Bundesärztekammer, Karsten Vilmar, has also complained, in a radio interview, that embryo research is imperative and that scientific freedom is endangered by laws that are too restrictive.

Infertility is a growing problem in West Germany, but Health Minister Süßmuth wants to ensure that *in vitro* fertilization (IVF) does not become a "tool for family planning". Couples who want but cannot have children comprise 10 to 15 per cent of all married couples in West Germany. Süßmuth suggests more intensive counselling as well as research into the causes of infertility, with IVF to be used only as a last resort.

The Ministry for Youth, Families, Women and Health will present a proposal for a law on surrogate motherhood later this year. The Justice Ministry is responsible for proposing a law covering embryo research.

Steven Dickman

DNA fingerprinting to be used for British immigrants?

London

THE technique of DNA fingerprinting based on restriction fragment length polymorphisms (RFLP) appears to have created a dilemma for the British Home Office, which is considering whether to use the tests for assessing applications for immigration to Britain from the Indian sub-continent.

Routine use of DNA fingerprinting in immigration cases would mean that more cases were approved and that £7 million a year in administration costs would be saved, according to a report being considered by the Home Office. In British law, resident immigrants are entitled to ask for resident status for certain relatives.

Two weeks ago, the Home Office minister responsible for immigration, Mr Timothy Renton, was quoted as agreeing that the test can categorically say whether or not "a person is related to the father or the mother... but not how they are related". But that is "wholly wrong", says Dr Alec Jeffreys, of Leicester University, who invented the test and helped the company Cellmark Diagnostics carry out and report on the pilot project commissioned by the Home Office.

Cellmark Diagnostics, in which the chemical manufacturer ICI plc owns a controlling interest, has exclusive rights to commercialize the DNA fingerprinting test, and has considerable experience of its application to immigration cases. In its first 8 months of business, about 70 per cent of more than 2,000 tests have been for immigration cases, which on average involve fingerprinting blood samples from about five individuals.

Most cases involve a male resident in Britain applying for permission for his wife and children in Pakistan or Bangladesh to join him. According to Mr Ron Yaxley of Cellmark, the claimed relationships turn out to be accurate in about 95 per cent of cases, whereas only about 50 per cent of claims are validated by the traditional means of interviewing.

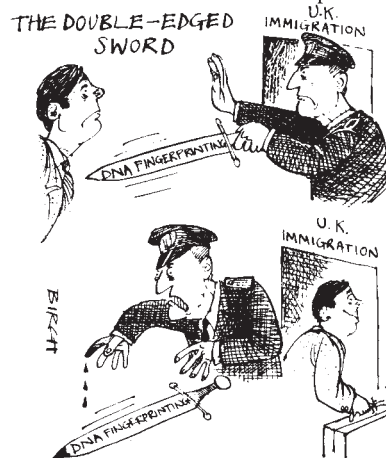
The pilot study of 40 immigration cases commissioned by the Home Office also shows that most claims are valid. The Home Office did query some of the results, says Yaxley, but they survived checking.

On the question of whether the test can always determine precise relationships, Yaxley and Jeffreys say that in the overwhelming majority of cases there is no problem, a point conceded by a Home Office spokesman last week in apparent contradiction of the minister's earlier statement.

Sorting out relationships is at its most difficult when, for example, some of the children claimed by a 'father' are in fact

children of either his brother or his sister-in-law. It is easy to spot that they are not his, but not so easy to determine to which of his two relations they belong. The problem is resolved by new 'single locus probes', which complement those that determine variations in several chromosomal loci simultaneously.

Care must also be taken, says Jeffreys, to allow for new mutations — the extreme variability of certain DNA loci that accounts for the variations between individuals that are detected as DNA fingerprints is the result of a very high rate of mutation at the loci, so that a difference between a child and its two claimed parents may be the result of a new mutation rather than a false claim of parentage.



The adoption of routine DNA fingerprinting by immigration authorities has been urged by immigrants' organizations,

which believe it should speed up decisions and prevent many applications being wrongfully dismissed. The current system depends largely on interviews.

Dr David Pearl, who is a part-time adjudicator of disputed decisions, has pointed out in a recent article in *The Law Society's Gazette* that claims falsely made in other circumstances, perhaps for tax purposes, can come home to roost in subsequent immigration hearings. In particular, men falsely claiming wives and children who subsequently marry and sire children when visiting the sub-continent inevitably "get themselves into a horrible mess, and the whole application is almost bound to fail".

A Home Office decision on the adoption of routine testing is eagerly awaited by immigrants' organizations, who fear that it may be delayed by the political problems of approving a measure that increases immigration. Because the Home Office and the criminal courts have already approved the test in forensic cases, it is most unlikely that the test itself will be rejected for immigration cases. Moreover, for the purposes of the pilot study, the Home Office has already devised procedures to prevent the supply of false blood samples — blood has to be taken by an approved doctor under supervision in the British Embassy in Dacca or Islamabad, or by a consultant haematologist at a general hospital in Britain.

Whether the cost of testing will count against it in the Home Office's eyes remains to be seen. At an average testing cost of about £500 a case, Cellmark Diagnostics estimates an annual saving to the Home Office of £7 million, but Home Office costings are likely to differ.

Peter Newmark

Mystery of bird deaths in Assam

New Delhi

AN eight-year study by Indian zoologists has failed to establish why birds commit suicide year after year at the small village of Jatinga in the northeastern state of Assam.

Attracted by the lights, birds converge on Jatinga at night and on landing become immobile, stop feeding and starve. They neither resist capture nor try to fly away.

The peak suicide season is in September and October, when the Indian monsoon tapers off, and the place is always the same: a one-kilometre stretch between the railway station and the health centre. Knowing that birds are lured by lamps, the villagers light up the sky with Petromax lamps on tall bamboo poles to attract more birds, which are then barbecued and eaten.

The mysterious phenomenon has been occurring since 1905, according to Dr Sudin Sengupta of the Zoological Survey of India (ZSI), who launched a systematic scientific study in 1979. Some 25 expeditions to

Jatinga have so far taken place but the Minister of Environment and Forests, Mr Z.R. Ansari, recently told parliament that "ZSI has not come to any conclusion as to the cause of the phenomenon."

Meanwhile, the state government of Assam has taken two steps to save the birds. Tall towers with powerful light sources have been erected to divert the birds from the light sources of Jatinga, and bird-watcher clubs have been formed to protect the dazed birds from villagers.

According to Sengupta, as many as 500 birds die every night during the peak season. Some 36 species of birds have been identified among the dead, including migratory birds. The studies have shown that the phenomenon, which takes place on dark rainy nights when winds are strong and clouds are dense, may be related to observed changes in geomagnetic and electric fields that apparently affect the bird's sense of orientation.

K.S. Jayaraman

Livermore and UC labs join forces in mass spectroscopy

Livermore, California

LAWRENCE Livermore National Laboratory (LLNL) and the University of California (UC) have become research partners in an accelerator mass spectroscopy (AMS) facility to be part of LLNL's new Van der Graaf tandem accelerator. The collaboration, welcome to LLNL as a chance to improve relations with UC and diminish the laboratory's image as a bomb factory, was initiated by UC archaeologists who want to use AMS for highly sensitive carbon-dating, without having to wait in line for one of the 20 or so heavily subscribed AMS facilities elsewhere.

Such collaborative projects are likely to become more common under the new UC contract for management of the Lawrence Livermore, Los Alamos and Lawrence Berkeley national laboratories, signed last September. The contract calls for an increase in scientific collaboration between the laboratories and UC, funded in part by the management fees collected by the university.

For its contribution of \$250,000 — for construction of the AMS ion source and detector system — UC will be one of 5 partners in the accelerator, sharing the 4,000 hours of scheduled annual operating time equally with 3 LLNL departments and Sandia Laboratories in Albuquerque.

The sensitivity of AMS carbon dating derives from its ability to count every atom of an isotope, after accelerating up to 5 per cent of the atoms in the sample, and deflecting their paths based on atomic mass. The technique is up to 6 orders of magnitude more sensitive than traditional carbon dating based on radioactive decay.

LLNL scientists also plan to use AMS for composition analysis of proposed nuclear-waste disposal sites, as well as biomedical applications. AMS may also prove useful for nuclear weapons test diagnostics, but it is unlikely that the laboratory will do weapons experiments on the UC-funded beam line.

The AMS will be one of 19 beam lines and experiment stations of the new Van der Graaf tandem accelerator that will be used in a variety of research areas, including radioactive ion beams, nuclear spectroscopy and materials science. The entire facility cost \$4 million, a fraction of the potential price, thanks to innovative design, and bartering by Jay Davis, head of LLNL's nuclear physics division. It was built entirely from used parts, including the accelerator itself, acquired from the University of Washington, and the magnets for the AMS system, which were a gift from the Stanford Linear Accelerator Center.

Marcia Barinaga

Note worth the plastic it's printed on

Sydney

THE world's first banknote printed on clear plastic has been issued in Australia. The \$10 note includes anti-forgery devices designed to beat even the most sophisticated photographic copying. The print technology has been developed by Australia's Reserve Bank and the Commonwealth Scientific and Industrial Research Organisation (CSIRO) at a cost of A\$20 million.

The note has come up shining after immersion in beer, petrol, grease and tomato sauce, being put through a washing machine, boiled and buried for nine months. The plastic note has much the same thickness and feel as its paper predecessor, though perhaps a little greasier. The Reserve Bank is being cagey about the type of plastic used. The main breakthrough necessary to make the new note was getting the ink to adhere to the plastic substrate through thick and thin.

The note's main security feature is the eye-catching, almost gaudy, rainbow-coloured optically variable device (OVD), an aluminium diffraction grating image of Captain Cook, whose colours change with the angle of viewing. The layer of aluminium is so thin that it cannot be felt with

the fingers. Another anti-photocopying device is the transparent window surrounding the OVD. Other innovations include micro-printing which is sharp on a scale smaller than easily seen with the naked eye.

In several South American countries notes are printed on a substrate of compressed plastic fibrils having much the same structure as paper, but none has ever been printed on clear plastic sheet before.

One side of the new note shows the *HMS Supply*, one of the First Fleet, and on the other side Aboriginal motifs including a boy wearing body paint, an ancient rock painting of a female figure, hand stencils and a ceremonial Morning Star Pole.

The commemorative note — issued in 1988 — to commemorate the 200th anniversary of European settlement in Australia, will be a test for the new technology. The new notes are more expensive to produce than their paper predecessors, but if they stand up to the rigours of use for longer than the average eight months of paper ten dollar notes then it won't be long until Australia's paper money is all plastic. The anti-forgery technology should soon appear in other high security applications such as passports.

Charles Morgan

SDI hails success in space test

Washington

THE Strategic Defense Initiative Organization (SDIO) last week successfully orchestrated a \$250-million space mission to study detection and tracking of ballistic missiles.

SDIO described the mission, called Delta 181, as "one of the most complex unmanned Earth orbit space missions ever attempted by the United States". The National Aeronautics and Space Administration (NASA) on 8 February used a Delta rocket to launch the three-tonne payload into orbit for SDIO from the Kennedy Space Center at Cape Canaveral, Florida.

Tracking objects from space is a critical requirement for a ballistic missile defence. The Delta 181 payload carried passive sensors, including ultraviolet and infrared imagers and spectrometers, as well as active sensors such as a pulsed laser radar (lidar), a coherent Doppler lidar and a microwave radar.

Mission director Major Green said one instrument, an infrared imager, failed during the 12-hour data collection period, but other instruments would provide much of the same data.

The targets used in the exercise were 14 different test objects released by the Delta 181 once it had reached its orbit. SDIO is particularly interested in how well targets can be tracked against different backgrounds, as well as through the Earth's limb, just above the horizon.

Another experiment tracked boosting missiles to discriminate between the target and the hot gases from the rocket's engine, for which purpose the Delta 181 carried four targets with solid rocket motors. SDIO launched a three-stage Strye 11 sounding rocket from Hawaii that was tracked from orbit.

Approximately 10¹¹ bits of data were collected by the Delta 181 mission, which is to be transmitted to ground stations over a ten-day period. Mission organizers will not explain why so much time is needed to transmit that much data, which could be handled by the Tracking and Data Relay System operated by the Goddard Space Flight Center in something like 20 minutes. Some have speculated that the extra time in space may be used to conduct experiments not yet announced, such as tracking other ground-launched targets.

Even critics of SDIO's plans to build a ballistic missile defence acknowledge that a mission such as Delta 181 is a necessary precursor to building a comprehensive system, but say that success does not prove that the concept of SDI is valid.

Joseph Palca

Objectivity in science

SIR—Despite some unfortunate, inaccurate and even outrageous remarks about philosophers of sciences the community of science is greatly indebted to Theocharis and Psimopoulos (*Nature* **329**, 595; 1987). They have raised questions that call for urgent and honest answers. Not least, they have brought into the open the problems of objectivity and truth in science and, perhaps, reminded us of what we are in danger of losing. But some cautionary notes need to be sounded.

In suggesting that the practice of science today is affected (adversely) by the prevailing whims and fashions of 'philosophy', are they not in mortal danger of selling the pass to their enemies? Is this not precisely the point of the cultural relativists whom they oppose? According to these "betrayers of the truth", science is always conditioned by cultural values and it cannot be said to deal with unchanging objective realities. The present difficulties vexing the Royal Society, for instance, are a standing demonstration of the critical dependence of science on socially acceptable values. Yet this is to fudge the distinction between public (including government) attitudes to science on the one hand and the attitudes of the scientists on the other. It by no means follows that the one is reproduced in the other. The opposite may well be true.

If that is a distinction between the piper and he who calls the tune, there is another to be drawn between music-maker and critic. Few would deny the latter the right to exist, but it is to be hoped he or she is not tone-deaf. When criticism of science is concerned, however, tone-deafness (or its equivalent) is all too prevalent. The public understanding of science is often formed by those with no experience of doing science, and great harm is done by their guileless and simplistic pronouncements.

A further distinction may further clarify the issues. It is that between scientific *method* and scientific *intention*. One reason why scientists are so embarrassingly coy about their methods is that these vary widely over the whole field of science, as witness the different approaches to (say) evolutionary theory and antibiotic therapy. (Another reason is simply that the methods have become second nature and are rarely defined in words.) But this is not to imply a diversity of intention, for there always is a determination to wrest secrets from nature, whether or not this leads to 'practical' applications. Indeed, it could be argued that it is this very uniformity of intention that binds the scientific community together.

It is remarkable that Theocharis and Psimopoulos make little mention of the history of science (apart from a passing reference to an unspecified golden age).

For it is studies in this area that throw into sharp relief yet another distinction, namely that between *actual* and *public* motivations. Over the past 20 years or so, an immense amount of painstaking scholarship has been devoted to private papers, diaries and other manuscript sources, and has produced overwhelming evidence that the practice of science is very culturally dependent. Yet even that overlooks a further distinction which is probably more important than any of the others. It is the difference between approaches to science, which may be conditioned by our culture, and the empirical results determined by nature.

Whether the buzz-words 'objectivity', 'truth' or 'value-transcendence' are actually used or not, common observation suggests that few scientists doubt that there is something 'out there' to be examined, whatever their own circumstances may be. That pragmatic belief today is successor to a deeply held Christian conviction about the created world that attended the rise of modern science. Donald MacKay has argued that "the ideal of objectivity . . . is the main spring behind the success of science, and finds its status reinforced by that success"¹. Such a view is entirely compatible with historical research about the ways in which scientists formulate programmes, examine and select data and interpret their results. It is not that we have to choose one or the other: culturally conditioned methodology or an objective world outside. We need both. Only then can we appreciate the essential humanness of science, but only then can we find any hope of escape from the miasmatic swamps of sociological relativism from which the image of science so urgently needs to be rescued. To set the record straight is a noble agenda for history of science for the rest of this century.

COLIN A. RUSSELL

*Department of History of Science
and Technology,
The Open University, Walton Hall,
Milton Keynes MK7 6AA, UK*

1. MacKay, D. *Objective Knowledge* (ed. Helm, P.) 43–57 (Inter-Varsity, Leicester, 1987).

• This correspondence is now closed. —
Editor, *Nature*.

Editorial policy

SIR—In his article on evidence for an influx of small comets into the Solar System, Paul Feldman (*Nature* **330**, 518; 1987) also comments on the editorial policy of *Geophysical Research Letters*. He asks whether the papers of Frank *et al.*¹, which inspired the *Nature* article² whose findings Feldman discusses, should have been published at all. He also disagrees with the policy of *Geophysical*

Research Letters that gave Frank and his co-authors the opportunity to respond to each criticism that their papers drew. I am pleased to outline the editorial policy that governed the publication of the original papers by Frank *et al.* and of the ensuing comments and replies.

Geophysical Research Letters seeks to publish papers that describe interesting, forefront science. These papers occasionally challenge conventional wisdom and thereby engender controversy. Written at the turbulent interface between ignorance and knowledge, such papers are frequently wrong. However, not all controversial ideas are wrong. The importance for scientific progress of the occasional new idea that proves correct is out of all proportion to the number of such ideas. Because it is not possible to tell in advance which new idea is correct, it is best to get them into the open literature where they can be discussed, attacked, tested or supported as the will of the community and the soundness of the idea dictate. This is the way science has advanced, and no better way has been demonstrated.

Exposure of conflict is a proper function for a letters journal that publishes new and interesting research results. Comments and replies are an important part of *Geophysical Research Letters*. They provide a forum for discussion and testing of new ideas. Rather than restricting such debate to private exchanges between authors and referees, the publication of comments and replies allows the broad scientific community to hear both sides of an issue and form an independent judgment.

Finally, I would claim that the publication by *Nature* of the paper by Donahue *et al.* is, in itself, a vindication of the original decision to publish the papers of Frank *et al.* Their papers and the ensuing debate will, no doubt, inspire other useful work. And, although Feldman is perhaps annoyed by the protestations in the replies of Frank *et al.* that their critics have not shown conclusively that the claimed comets do not exist (in the sense that claims are made that the non-existence of the Loch Ness monster has not been conclusively proved), it seems clear from his statements that Feldman himself has not been persuaded. I don't believe many others have been either.

A. J. DESSLER
(Editor)

*Geophysical Research Letters,
Department of Space Physics and
Astronomy, Rice University,
Houston, Texas 77251-1892, USA*

1. Frank, L.A. *et al.* *Geophys. Res. Lett.* **13**, 303–306 (1986); *Geophys. Res. Lett.* **13**, 307–310 (1986).

2. Donahue, T.M. *et al.* *Nature* **330**, 548–550 (1987).

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only. □

Beating the quantum limits (cont'd)

Heisenberg's Uncertainty Principle is for many an irksome constraint on the freedom to make measurements accurately. Can the constraint be overturned?

Is it possible to evade the restrictions of Heisenberg's uncertainty principle by the clever design of measuring equipment? Plainly, the answer to a question such as this would be of great practical importance in, say, the information processing industry. The trend towards ever-smaller processing devices, for example, will ultimately be limited by quantum mechanics, or by the fidelity with which a device put into a specific state so as to represent a bit of information will persist in that state despite unavoidable interactions with external systems.

Historically, the search for stratagems to beat the quantum limit has, in practice, been driven by those looking for gravitational waves, and who are naturally anxious to beat the quantum limit in their design of the masses which, supposedly, would be set vibrating by interaction with gravitational waves. Inevitably, the argument centres around the question of precisely what is meant by the process of making a quantum measurement.

As is customary in these connections, the vigorous arguments there have been over the past several years have entailed the splitting of several recondite hairs. Until recently, it has seemed that those who hold that the "standard quantum limit" (SQL) is inescapable were in the ascendant. But now Masanao Ozawa of Nagoya University has put a cat among the pigeons by specifying a quantum system in which, he says, it is possible to do better than SQL (*Phys. Rev. Lett.* **54**, 2465; 1988).

The conventional view has been put well by Carlton Caves of the California Institute of Technology (*Phys. Rev. Lett.* **54**, 2465; 1985) for the case of what must be the simplest quantum measurement of all, that of the position of a free particle at two instants separated in time by an interval τ . The strength of Caves's argument stems in part from his ready acceptance of an earlier assertion by Horace P. Yuen (*Phys. Rev. Lett.* **51**, 719; 1983) that there is indeed a flaw in the standard textbook derivation of SQL.

That argument is as follows. Suppose the measurement of the position of the particle at time t_0 is uncertain by an amount Δx , so that the corresponding uncertainty of the momentum (p) is that given by the Heisenberg relation as Δp , which must be $\geq \hbar/2\Delta x$. The uncertainty of the position of the particle after an inter-

val τ is then a function of two parts — the uncertainty of its position at the beginning of the interval and the extent to which the position of the particle is further spread by the spread of the possible momentum (or velocity) after the first measurement. Combining the two components as statistical variances, the textbooks show that $(\Delta x)^2$ after an interval τ is $\geq \hbar^2\tau^2/m$, where m is the mass of the particle.

The flaw in this argument, pointed out by Yuen and now, apparently, generally accepted, is that there is more to say about the evolution of the quantum system during the interval τ . In terms of the operators representing position and momentum, $x(t) = x(t_0) + p(t_0)t/m$. In suitable circumstances, the outcome at the end of the interval is a correlation term which may be negative and which, when subtracted from the variances due to the initial uncertainties, will help to beat the quantum limit. Everything depends on the state in which the initial measurement leaves the particle.

Caves's counterargument, reminiscent of the things that people were saying at Copenhagen in the 1920s, is that the old textbook argument is deficient in its neglect of the measuring equipment. Specifically, he defines the imperfect "resolution" of the equipment as a quantity σ whose square must be added to the variance of the position of the free particle to yield the uncertainty of the second measurement. Whether or not there are states of the particle, called "contractive states" by Yuen, the SQL is a simple consequence of Heisenberg's Uncertainty Principle in Caves's view. One interesting caveat in the argument is the assumption that the coupling between the measuring apparatus and the system observed must be linear. In 1985, Caves was guessing that those seeking violations of SQL should see what non-linear couplings have to offer.

What Ozawa has now done is to pull apart Caves's definition of "resolution", which he says may mean either the uncertainty of the result of a measurement, the uncertainty in the position of the particle after the measurement and the uncertainty in the measuring apparatus immediately before it. He defines instead the "precision" of the measurement, which is the uncertainty that arises in the measurement of the position of a particle known to be in an eigenstate of the position oper-

ator (and whose momentum is thus entirely uncertain) and the "resolution", which is the difference between the result of a measurement and the position of the free particle just after it. The question is whether the precision can ever be less than the resolution, as thus defined.

Naturally, Ozawa's conclusion is that such a state of affairs is indeed attainable. The crux of his argument is the construction of a solvable model to represent the interaction between the measuring equipment and the free particle whose position is to be measured, which has the virtue that (regarded as a quantum mechanical hamiltonian) it can be solved exactly. In reality, what he has done is to show that, in Yuen's language, contractive states exist and have exactly the properties predicted for them. On the face of things, it seems, the SQL can be beaten by this means to any desired extent.

Against the background of past arguments, it seems unlikely that this argument, neat though it is, will turn the conventional position on the SQL. One of the virtues of Ozawa's case is that the quantities arising in his calculations are indeed precisely defined and are related directly to quantities that can be measured. On the other hand, there are difficulties about the assumptions made about the way that noise associated with the measuring system would be allowed for as well as in the formality with which the yardstick itself is defined. Yet Ozawa's calculation will undoubtedly lift the spirits of those involved with the design of gravitational wave detectors; it will be interesting to see where this leads.

But, as the past half century has shown, this is not a field in which arguments come neatly to generally accepted conclusions. On this occasion, there may be continuing arguments about Ozawa's assumption that the coupling between the free particle and the measuring equipment is so strong that the unperturbed motions of the particle and the yardstick can be neglected. It is also far from obvious how the particular example on which the conclusion rests can be turned into a realistic measuring equipment that would allow those who design equipment to exploit this recipe for beating SQL. But none of this will damp enthusiasm for overturning what often seems an intolerable constraint on the freedom to design accurate measuring equipment.

John Maddox

Cognitive neurophysiology

The lifeblood of language

John C. Marshall

IT WOULD be useful, to put the matter mildly, if we could see what the human brain was doing when the brain was doing it. And it would be yet more valuable to be able to visualize the neuronal substrate of complex abilities that can be fractionated into simpler, independently assessed stages. For the first time, Marcus Raichle's team has now successfully wedded an advanced brain-imaging technique for measuring blood flow within defined cortical regions to an information-processing model of the steps involved in some basic language functions. This work, reported by Petersen *et al.* on page 585 of this issue¹, shows how cognitive theory and current medical technology can converge on a unified psychobiology of lexical processing.

The story begins somewhat obliquely. In 1890, Roy and Sherrington observed² that the onset of generalized epileptic seizures, provoked by synchronized hyperactivity of nerve cells, was rapidly followed by substantial swelling of the brain. This swelling, they conjectured, was due to increased blood-flow. And they were accordingly led to postulate the existence of "an automatic mechanism by which the blood supply of any part of the cerebral tissue is varied in accordance with the activity of the chemical changes which underlie the functional action of that part."

The basic hypothesis turned out to be correct, and by the 1960s, technological developments allowed the functional landscape of the human brain to be plotted *in vivo* by measurement of regional cerebral blood flow (rCBF). A radioactive isotope, xenon-133, that attaches to red blood cells was introduced into one of the main arteries that supply the brain's nutritional requirements; a gamma-camera could then monitor simultaneous radioactivity counts from different brain regions. And these counts could in turn be transformed into measurements of rCBF that reflected the differential activation of brain areas when the conscious patient (or normal volunteer) is actually engaged in a particular mental task³.

If the data from the gamma-camera are further processed in a computer, then images can be displayed on a colour monitor in which different hues are assigned to different levels of blood-flow. David Ingvar, one of the pioneers of rCBF methodology, has therefore termed such flow-charts "cerebral ideograms" of mental activity⁴.

In general, early studies of speech and language processes using the xenon-133

technique supported the classical views of localization of function that had been derived from investigation of the effects of focal damage to the left hemisphere of the brain. For example: in neurological patients without focal tissue damage, speaking aloud activates auditory cortex,



Watching the working brain.

the primary (rolandic) mouth-tongue-larynx area and the supplementary motor area; silent reading activates the visual-association area, the frontal eye field, supplementary motor cortex and the foot of the third frontal convolution (Broca's area)⁵. There were, however, some surprises for traditional neuropsychology; most notably, increased flow in right-hemisphere areas that, when damaged, do not typically provoke aphasic impairment⁶. This finding should alert us to the possibility that inhibition, as well as excitation, requires energy.

Such studies promised considerable insight into the neurophysiology of language processing but were subject to several technical limitations: poor spatial resolution of adjacent brain areas; imprecision of anatomical localization due, in part, to attenuation of radiation from structures deeper in the brain; and poor temporal resolution due to the slow clearance curve for ¹³³Xe. Equally critically, these investigations used multi-faceted tasks that 'lit up' very large regions of the cortical mantle.

By contrast, Raichle's team¹ has devised a set of interrelated behavioural and imaging techniques that are rapidly overcoming the inadequacies of previous work. Positron emission tomography (PET) can image rCBF by reading the spatio-temporal distribution of injected radionuclides with short half-lives

(oxygen-15, for example). As each individual radiation dose is low, repeated measurements can be taken from the same subject performing different tasks. Although the spatial resolution of adjacent point sources by PET is currently poor (1.0 cm at best), a single point source can be localized with excellent accuracy (5 mm or better).

Hence, if experiments are designed so that each new behavioural task adds a further processing requirement to the previous one, rCBF measurements from the simpler task can be subtracted to obtain a precise image of the successive stages of the more complex task. Raichle's group initially validated the subtraction method by imaging the retinotopic map within human visual cortex⁶; it is now extended¹ to help resolve a longstanding controversy about visual word recognition and reading aloud.

The first models of reading, proposed by nineteenth-century neurologists⁷, explicitly claimed that all reading for meaning was mediated by the cortical centre for auditory speech processing. Lesions that impaired the ability to read (and write) centred around Wernicke's area and the angular gyrus (in the left posterior temporo-parietal region), areas that are known to be crucial to the analysis and comprehension of speech. This model dominated the interpretation of both normal and pathological reading until the mid-1970s. LaBerge thus categorically states⁸ that "written material must be coded into phonological form to be comprehended." The model was challenged, however, by the discovery⁹ of patients who could read words aloud (albeit with numerous semantic para-

Welcome to Daedalus

DAEDALUS, who will now be contributing regularly to this section of *Nature* (see page 569 of this issue), is a senior member of the British company DREADCO, described as "the innovation factory" in the prospectus in which, last year, it asked its shareholders for extra funds. If readers wish to think of Daedalus as the one who first arranged the backing of venture capitalists for DREADCO's launching, they are free to do so.

As the weeks pass, it will be found that Daedalus epitomizes the boast in last year's prospectus that DREADCO is the only "proven high time-rate broad-spectrum ideas company in Western Europe", but the cultivation of a sense of irony will be found an advantage in understanding them. Because of the "rigorous corporate quality-control procedures" to which contributions from Daedalus will be subjected, the hazards of the formal refereeing process will, on these occasions only, be dispensed with. Unsolicited referees' reports will be forwarded to the author. □

lexias, such as reading 'sick' as 'ill') but could not read even the simplest non-words (such as 'nol', 'dup' or 'tud'). This pattern of performance suggested the existence of a reading route whereby the visual stimulus addressed semantic content directly, without internal phonological recoding.

Petersen *et al.*¹ use their ingenious 'subtraction' methodology to track the distinct stages of reading aloud in neurologically normal volunteers. To localize sensory processing, they contrast simple visual fixation with fixation plus (visual) word presentation (without any overt behavioural response). To localize motor programming and articulatory output, they contrast fixation and word presentation with the additional requirement to pronounce the presented word; this repetition task is then contrasted with a 'semantic association' condition in which the subject gives a (single-word) use for each stimulus word (for example, if 'cake' is presented, say 'eat'). The relevant subtractions of rCBF images show that none of these tasks with visual-word input activates Wernicke's area or the angular gyrus. The technique is sufficiently sensitive to detect increased flow in these areas, for they are significantly activated by the presentation of auditory word stimuli.

These results thus appear to offer direct confirmation of a reading route that does not involve obligatory phonological recoding of the visual stimulus before semantic access. Further experiments immediately suggest themselves. What, for example, will be the pattern of flow when reading non-words (that cannot be assigned a semantic interpretation)¹⁰? Can the technique throw light on the hypothesis that there is a subsidiary, semantically based reading system in the right hemisphere¹¹? If the (theoretical) flow-charts of cognitive neuropsychology¹⁰⁻¹² (see figure) can continue to be matched with the (literal) flow-charts of cognitive neurophysiology, there are very exciting times ahead. □

- Petersen, S.E., Fox, P.T., Posner, M.I., Mintun, M. & Raichle, M.E. *Nature* **331**, 585-589 (1988).
- Roy, C.S. & Sherrington, C.S. *J. Physiol., Lond.* **11**, 85-108, (1890).
- Risberg, J. & Ingvar, D.H. *Expl Brain Res.* **5**, 72-78 (1968).
- Ingvar, D.H. *L'Encephale* **3**, 5-33 (1977).
- Lassen, N.A. & Larsen, B. *Phonetica* **37**, 27-37 (1980).
- Fox, P.T. *et al. Nature* **323**, 806-809 (1986).
- Lichtheim, L. *Brain* **7**, 433-484 (1885).
- LaBerge, D. in *Language by Eye and by Ear* (eds Kavanagh, J.F. & Mattingly, I.G.) 241-248 (MIT, Cambridge, Massachusetts, 1972).
- Marshall, J.C. & Newcombe, F. *J. psycholing. Res.* **2**, 175-199 (1973).
- Patterson, K.E., Marshall, J.C. & Coltheart, M. (eds) *Surface Dyslexia* (Erlbaum, London, 1985).
- Coltheart, M., Patterson, K.E. & Marshall, J.C. (eds) *Deep Dyslexia* 2nd edn (Routledge & Kegan Paul, London, 1987).
- Cooper, H. & Cooper, P. *Heads or the Art of Phrenology* (London Phrenology Co., London, 1983).

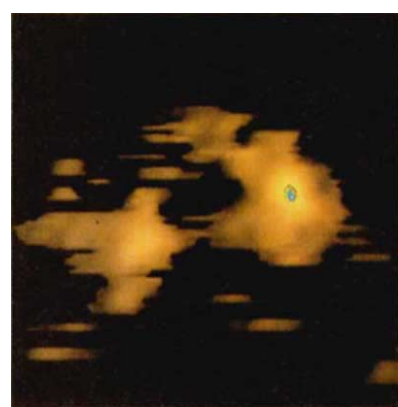
John C. Marshall is in the Neuropsychology Unit, part of the Neuroscience Group at the Radcliffe Infirmary, Oxford OX2 6HE, UK.

Cosmology

In search of young galaxies

Joseph Silk

THE discovery of forming galaxies would impose important restrictions on cosmological theories and ultimately lead to a unique scenario for the evolution of the large-scale structure. Most observed galaxies are mature systems that are already well down the evolutionary scale, and the quasars recently discovered at record-breaking redshifts (see the News and Views article by Peter Shaver¹) are too distant to reveal any structure that could be associated with forming galaxies. Observations and theory suggest that



The blue compact galaxy ESO400-G43. The surrounding cloud of neutral hydrogen and companion is shown in orange tone. An optical image of the galaxy is superimposed and is intensity coded in rainbow colours. Courtesy of the European Southern Observatory.

galaxy formation occurred at the epoch $z \approx 5$ or earlier. (Redshift, z , measures the shift in the wavelength of spectral lines in a distant object relative to that of a nearby galaxy, caused by the expansion of the Universe, and is a useful measure of the time that light has taken to travel from its source.) On page 589 of this issue², however, Bergvall and Jörsäter present evidence that the nearby dwarf galaxy ESO400-G43 is newly formed.

It is natural to assume, from the evidence of our own mature Milky Way and other nearby galaxies, that young galaxies must be remote in time as well as in space — that is, galaxies are not forming now. This is the basis of one theory of galaxy formation, originally proposed by Peebles and Dicke to account for globular clusters³. This theory has recently been resurrected⁴ under the guise of a model in which galaxies formed from primordial isothermal fluctuations (that is, fluctuations in matter density but not in radiation density). Isothermal fluctuations allow small seeds, of mass similar to that of dwarf galaxies, to develop at early epochs and are consistent with an open Universe

dominated by baryonic (ordinary) matter in which galaxy formation occurred at $z \approx 30$.

Young galaxies, however, are not necessarily present only in the very early history of the Universe. One theory that accords with this view and also appeals to hierarchical evolution of structure (in which dwarf galaxies form first, followed by larger and larger systems) but with non-linear evolution occurring only relatively recently, has been successful in accounting for several aspects of the large-scale structure of the Universe. The idea that the Universe is dominated by cold dark matter comes from inflationary cosmology, which prescribes a Universe at critical density with adiabatic fluctuations. These fluctuations involve both matter and radiation and are constrained by inflationary cosmology to have considerable power on large scales. This means that in this model, large galaxies are still forming at the present epoch. As is usual in cosmology, one suspects that the truth lies somewhere in between.

A galaxy in process of formation can be characterized in several ways. Perhaps it has no old stars. It may have an exceptionally high rate of star formation that could not be sustained for long and that can account for the observed luminosity. Or it may be predominantly gaseous, the gas being destined to form stars. Perhaps the galaxy is still gas rich but deficient in metals, a sure indicator of a primitive stage of chemical evolution. Or finally, the galaxy may be forming stars at high redshift, early in the evolution of the Universe. Candidate young galaxies generally satisfy at least two of these criteria.

In this issue, Bergvall and Jörsäter describe² an object known as ESO400-G43 (see figure) that fulfils several of these requirements. This blue, compact galaxy is forming stars prolifically, it is gas-rich, and it appears to contain no old stars. It also has a dark halo that Bergvall and Jörsäter traced by the 21-cm hydrogen radio emission and which extends to about 15 disk scale lengths. (This halo could be of use to cosmologists trying to derive mass limits for so-called heavy neutrinos. The inferred radius of the halo's core — beyond which the density tails off — and its inferred velocity dispersion give a lower limit of 80 eV for any single species of massive neutrino which would account for the dark-halo matter.) ESO400-G43 is also metal-poor, at about one-eighth of solar metallicity, and requires no old stars at all to account for the mass-to-light ratio of the disk.

This object should be warmly welcomed by advocates of the cold dark matter theory. A dark halo, an apparently young, chemically unevolved stellar component and a plentiful supply of gas are the ingredients expected for a young galaxy. Cold dark matter theory⁵ prescribes the continuous rapid formation of many dwarf galaxies from the typical density fluctuations; larger fluctuations collapse later. A few of these dwarfs are incorporated into luminous galaxies. It is necessary to introduce the idea of biasing into this theory to suppress the formation of an excessive number of young galaxies. According to our version⁶ of this idea, dwarf galaxy formation is interrupted by the onset of supernova-driven winds that strip out the gas. Perhaps during the fiery vigour of massive star formation, these objects would be recognizable as blue compact dwarfs. By the present epoch, however, one expects the Universe to be full of dwarf galaxies, pale remnants of blue, compact galaxies.

Occasionally, one might come across a rare, delayed fluctuation that recently collapsed. So Bergvall and Jörsäter² prefer

to attribute the blue and compact nature of ESO400-G43 to such an event, so that no underlying component of older stars can be expected. An alternative explanation is that a dwarf galaxy has encountered a cloud of weakly enriched gas that fuels the formation of new stars as it falls inwards. The massive halo provides a gravitational well capable of accreting intergalactic gas. The density of such gas can be estimated from observations of enriched gas in rich clusters, and is probably adequate in less dense regions of the Universe to fuel accretion. Future observations, especially in the infrared, will refine the search for an older stellar population, and so help discriminate between the two options of virgin birth or resurrection for the origin of such remarkable blue compact dwarfs. □

1. Shaver, P. *Nature* **330**, 426 (1987).
2. Bergvall, N. & Jörsäter, S. **331**, 589–591 (1988).
3. Peebles, P.J.E. & Dicke, R. *Astrophys. J.* **154**, 891–908 (1968).
4. Peebles, P.J.E. *Astrophys. J.* **315**, L73–L76 (1987).
5. Bardeen, J., Bond, J.R., Kaiser, N. & Szalay, A. *Astrophys. J.* **304**, 15–61 (1986).
6. Deikel, A. & Silk, J. *Astrophys. J.* **303**, 39–55 (1986).

Joseph Silk is in the Department of Astronomy, University of California at Berkeley, Berkeley, California 94720, USA.

Virology

A case of mistaken non-identity

Carel Mulder

IN 1985, when Daniel *et al.* isolated¹ a virus that resembled human immunodeficiency virus (HIV) from Rhesus macaques afflicted with simian AIDS — the monkey equivalent of human AIDS — the stage was set for new developments in understanding the range and relationships of immunodeficiency viruses. Unfortunately, the plot of the events that followed served largely to confuse spectators. The denouement appears in the exchange between Kestler, Daniel and collaborators and Essex and Kanki on pages 619 and 621 of this issue^{2,3}, respectively.

The Daniel *et al.* virus, now called SIV_{MAC} (but originally STLV-III_{MAC}), answered all of Koch's postulates as the causative agent of simian AIDS, a rare disease found in macaque colonies at the New England and Californian regional primate centres. (The more common immunodeficiency found in these macaque colonies, SAIDS, is caused by a non-lentivirus: Mason Pfizer monkey virus-like D-type retrovirus, also first isolated by Daniel, a virologist who discovered at least ten new virus groups.)

Thereafter, SIV_{MAC} was found by Essex and Kanki to react with sera from a sizeable proportion of African green monkeys as well as sera from healthy prostitutes in Senegal that did not react with HIV-1, a virus not known in West Africa at that time. Kanki and collabora-

tors subsequently reported⁴ the isolation from these Senegalese women of a virus they named HTLV-IV, and the isolation from captive and wild African green monkeys⁵, of a virus they called STLV-III_{AGM}. Shortly after, Clavel, Montagnier and co-workers reported the isolation of HIV-2 (originally named LAV-2) from HIV-1-negative West Africans with an AIDS-like disease. Nucleotide-sequence analysis⁶ of these new viruses showed that HIV-2 is much more closely related to SIV_{MAC} than to HIV-1.

For a time, there was no reason to be concerned about the identity of either STLV-III_{AGM} or HTLV-IV, but doubts began to arise, first as other groups were unable to reproduce the isolation of a virus from African green monkeys by the procedure published for STLV-III_{AGM}, and then as sequence data emerged.

With a change in procedure, Hayami's group in Tokyo has recently been able to isolate 19 SIV_{AGM} viruses (ref. 6; M. Hayami, personal communication) from captive and wild Ethiopian and Kenyan African green monkeys. Using this procedure, at least two other groups have now also obtained SIV_{AGM} isolates (R. Kurth *et al.* and M. Daniel, personal communications). The sequences of these SIV_{AGM} isolates appear to be very different from those of STLV-III_{AGM} but similar to each other with a variability comparable

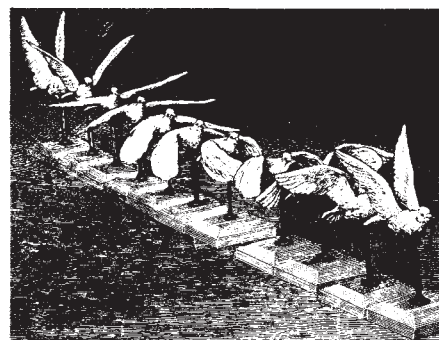
to that found among HIV-1, or HIV-2 isolates (M. Hayami and K. Ishikawa, personal communication).

Given the sequence variation among HIV-1 isolates that was already known at the time, it was very surprising when first Mullins's group^{7,8} and then Hahn's⁹ reported that six out of six STLV-III_{AGM} and three out of three HTLV-IV isolates had identical or nearly identical restriction-endonuclease cleavage patterns and nucleotide sequences. This is most unusual, both for lentiviruses in general and for HIV-like viruses in particular. Now, Daniel and collaborators, in their paper in this issue², find the explanation for this curious finding and solve the controversy on the origin of the nearly identical virus isolates. Whereas SIV_{MAC} isolated from different macaques have variable endonuclease patterns and sequences even when epidemiologically related, the restriction pattern of one of their isolates, SIV_{MAC251}, is identical to those of HTLV-IV and STLV-III_{AGM}; SIV_{MAC251} is also one of the two SIV_{MAC} strains provided to Kanki by Daniel¹⁰ for her seroepidemiological studies. It honours Essex and Kanki that they readily admit in their reply to Daniel's group on page 621 that all these questionable isolates should now be considered laboratory contaminations of their African green monkey and Senegalese human cells with SIV_{MAC}. Too seldom do researchers in this field retract data found to be erroneous.

The misidentification of the STLV-III_{AGM} and HTLV-IV isolates will not alter the results of their seroepidemiological surveys except that their Senegalese cohort should now be labelled HIV-2 seropositive. Both B. Hahn (personal communication) and Essex's group have now isolated genuine HIV-2 from the cohort of healthy Senegalese prostitutes, and Essex's group has also isolated HIV-2 from HIV-1-negative Senegalese AIDS patients (Marlink, personal communication). To minimize the confusion raised by

100 years ago

MECHANISM OF THE FLIGHT OF BIRDS



Bronze figures representing eleven successive positions in the stroke of a pigeon's wing achieved by the use of photochronography. From *Nature* **37**, 372; 16 February 1888.

the mislabelling of the SIV_{MAC} reisolations, I hope that all genuine West African-type human viruses will, henceforth, be labelled HIV-2 and all simian viruses SIV (SIV_{MAC}, SIV_{AGM} and so on) and that the terms HTLV-IV and STLV-III_{AGM} be reserved exclusively for the laboratory contaminants.

This episode should serve as a strong warning for all virologists working with multiple isolates to check any new isolates against viruses present in the laboratory. I am aware, or have been told, of at least five instances in other laboratories in the

United States and Europe where non-infected cell cultures became infected with HIV-1 in the same containment hood. □

1. Daniel, M.D. *et al. Science* **228**, 1201–1204 (1985).
2. Kestler, H.W. III *et al. Nature* **331**, 619–621 (1988).
3. Essex, M. & Kanki, P. *Nature* **331**, 621–622 (1988).
4. Kanki, P.J. *et al. Science* **232**, 238–243 (1986).
5. Kanki, P.J. *et al. Science* **230**, 951–954 (1985).
6. Chakrabarti, L. *et al. Nature* **328**, 543–547 (1987).
7. Kornfeld, H. *et al. Nature* **326**, 610–613 (1987).
8. Hirsch, V. *et al. Cell* **49**, 307–319 (1987).
9. Hahn, B. *et al. Nature* **330**, 184–188 (1987).
10. Desrosiers, R. *et al. Nature* **327**, 107 (1987).

Carel Mulder is Professor of Pharmacology, Molecular Genetics and Microbiology at the University of Massachusetts Medical School, Worcester, Massachusetts 01655, USA.

Species preservation

The CITES conservation circus

N. Mrosovsky

THE convention on international trade in endangered species of wild fauna and flora (CITES) has noble aims: to help save species endangered by international trade through the recognition that success demands not only cutting off the supply by clamping down on poachers, but also limiting the demand by making imports illegal in consumer countries. At biennial meetings, representatives from more than 95 countries can list species on the appendices to the convention. Listing establishes either a virtually total prohibition on export and import (appendix I) or a monitoring of trade by a permitting system (appendix II). Unfortunately, decisions taken at the most recent CITES meeting in Ottawa last July are not encouraging.

There are a great many endangered species, but not all of them are endangered by trade. According to criteria adopted at Berne in 1976 for placing species in the appendices, both the biological status and the extent to which the species is threatened by trade must be taken into account. Information provided should preferably include scientific reports on population size and range, or at least some indication of the potential causes of extinction. On the trade side, requirements depend on in which appendix the species is being listed. Even for appendix II listing (monitoring), evidence that actual or potential trade constitutes a threat to survival is needed.

There are obvious reasons for limiting a convention on trade to species endangered by trade. Even this is a huge task: so many different countries, animals and plants are involved; there is the danger of re-export of products in different forms; and the problem of distinguishing endangered species from others that may look similar. To achieve its aims, CITES must be practical and credible. If it becomes too complex it will be unenforceable, and if it seeks to stop trade in animals that are not endangered by trade it may lose support.

The case of the poison-arrow frogs of

the genus *Dendrobates* illustrates the problems. These animals are small, generally brightly coloured amphibians found in tropical central and South America. Most species have probably not been used by Indians for poisoning arrows, although they do contain toxins. There are about 50 species in the genus, some arboreal and hard to study despite being conspicuous.

At the most recent CITES meeting, the representative from Surinam proposed



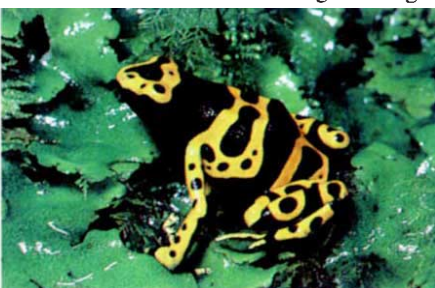
Over-protection — is the monitoring of trade in poison-arrow frogs such as *Dendrobates pumilio* (left) and *D. leucomelas* (right) really necessary? Photographs: Bruce Coleman.

that the whole *Dendrobates* genus be placed in appendix II. Few population data exist, nor is there any evidence that *Dendrobates* spp. are in danger of extinction. There is some demand in Europe and elsewhere for these colourful frogs to adorn terraria, but there is no evidence that this trade is affecting wild populations. *Dendrobates* are quite easy to breed in captivity and some of the pets in Europe come from such stocks.

Not only did the proposal to list *Dendrobates* on appendix II fail to meet the Berne criteria, but other information presented at the meeting strongly indicated that these frogs should not be listed. This information was compiled by the IUCN Trade Specialist Group and was based largely on the work of Charles Myers (American Museum of Natural History), one of the few people to have studied poison-arrow frogs extensively in the wild. Among the cogent points are: (1) the species appearing most often in trade are

mainly those with the largest ranges, and are often among the most abundant vertebrates in their habitats; (2) apparent small ranges for some species may simply be the result of inadequate study of other areas; (3) there is no evidence that these frogs are dependent on primary forest — they occur in mixed second growth and in banana plantations, for example; and (4) collection of large numbers of frogs does not seem to depress local populations. Sometimes they seem even more abundant after collection. Amie Brautigam, compiler of the IUCN report, notes that “Myers reports that during a 4-year study on *D. histrionicus* during which he collected 7,600 specimens from one population for scientific purposes, the population actually increased rather than decreased”. Taken together, there is no evidence that *Dendrobates* are endangered at present by trade — or by anything else. Yet the proposal from Surinam to put them on appendix II was passed, largely with the support of the South American countries who may have voted more in a spirit of regional solidarity than on the basis of scientific evidence.

The listing of such species may seem like a victory or at least a symbolic gesture for conservation, but is it really helpful or is it counterproductive? Burdening administrators and customs agents might



distract effort from species that are incontestably being threatened by trade. Indiscriminate listing of species and regional block voting, especially in the scientific committee, could turn CITES into a circus for political exhibitionism. Endangered species deserve the best science, not political posturing. Accepting flawed proposals actually discourages the research into the predicament of a species and the field investigations that should precede an adequate proposal. The poison-arrow frogs are only one example of the listing of species on the appendices without scientific justification. At its best, CITES is a splendid example of cooperation between rich and poor countries, but if it does not place the utmost importance on scientific considerations, it may survive for even less time than some of the species that need its help. □

N. Mrosovsky is in the Department of Zoology at the University of Toronto, Toronto, Ontario, Canada M5S 1A1.

Inorganic chemistry

Trapped electron pairs

Peter P. Edwards

THE ionic bond is the strong, all-directional electrostatic interaction between charged species such as a cation and an anion — for example K^+Cl^- . The simplest anion is the electron, and on page 599 of this issue¹, James L. Dye and co-workers report the preparation, structure and electronic properties of one of a remarkable new class of ionic salts called electrides. In these crystalline solids, stoichiometric amounts of electrons serve as the anion to a complexed, or cryptated, potassium cation, $K^+(cryptand[2.2.2])$. Importantly, the properties of this electride are consistent with the presence of strong electron-electron interactions leading to a 'singlet' (spin-paired) ground state, reminiscent of spin pairing in fluid metal-ammonia solutions.

The preparation and study of solid electride salts has its origins in investigations of solutions of sodium and potassium metals in anhydrous liquid ammonia, which were first examined by Sir Humphry Davy in 1807 (see ref. 2). Various metals (M) dissolve in liquid ammonia to produce intensely coloured blue and bronze solutions in which solvated electrons, solvated cations and various cluster or aggregate species

co-exist in equilibrium. Kraus, in 1908, identified³ that a valence electron, detached from its parent atom via the dissolution process $M(s) \rightarrow M^+ + e^-$ (s represents solid) exists in solution as a negatively charged entity surrounded by "an envelope of solvent molecules" (Fig. 1). Even in moderately concentrated metal-ammonia solutions, the possibility of interactions between solvated electrons clearly arises. Magnetic susceptibility measurements, first carried out⁴ by Taylor

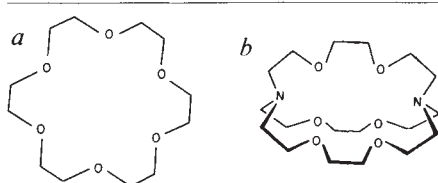


Fig. 2 Molecular structures of (a) 18-crown-6 and (b) cryptand [2.2.2].

and Lewis in 1925, reveal the presence of high concentrations of diamagnetic, electron-spin-paired species. Although the precise structure of this state is still controversial, two possible contenders are the dielectron species, e_2^- , (possibly in association with a cation) and the triple-ion species, $e^-M^+e^-$.

In contrast, magnetic optical and transport measurements on solutions of metals in amines and ethers reveal that the main diamagnetic species in these solutions is the distinct alkali anion, M^- , formed via⁵ $2M(s) \rightarrow M^+ + M^-$. In the case of the sodium anion, there is strong experimental evidence that Na^- interacts only very weakly with its environment. For the heavier anions, such as Cs^- , some dissociation can also occur to yield cations and solvated electrons: $M^- \rightarrow M^+ + 2e^-$.

A significant experimental problem in the study of the alkali anions was the low solubility of these metals in amines and ethers. The situation was drastically changed in the early 1970s by the seminal work³ of Dye and co-workers at Michigan State University. The experimental solution to these problems is provided by the complexation of the alkali cation M^+ by a crown ether or cryptand molecule (L) (Fig. 2). The complexation reaction $M^+ + L \rightarrow M^+L$ not only greatly enhances the metal solubility in these solutions, but also provides enough stabilization of M^- for 'alkalide' salts M^+LM^- (s) to be formed that are thermodynamically stable with respect to dissociation into the parent metal and the free complexant. The key to the formation of alkalide salts is the stabilization of the alkali cation by a large non-reducible cyclic or bicyclic complexant.

The structural features and physical properties of alkalide salts (with the exception of the salt of Li^- , the lithide ion) are now well established. It is clear that these anions have two electrons in the outer s-orbital and that they are the largest monoatomic anions in the periodic table. Even Na^- , the smallest alkalide, is larger than I^- , and the largest, Cs^- , has an effective diameter of about 7 Å.

Dye realized from the outset that it might also be possible to form salts in which the anions consist of trapped electrons, rather than M^- , because the addition of excess complexing agent to a solution which contains M^+L and M^- can also bring about the dissociation reaction $M^+ + L \rightarrow M^+L + 2e^-$.

The search for crystalline 'electrides' ($M^+L \cdot e^-$) has been a long one, plagued by difficult problems in sample instability and irreversible decomposition⁶. Eventually, in 1983, the first crystalline electride $Cs^+(18-crown-6) \cdot e^-$ was identified and its structure determined. In this material, single electrons are localized at individual lattice sites and exhibit weak antiferromagnetism. The new material $K^+(cryptand[2.2.2]) \cdot e^-$ is remarkable in that it has a plasma-like optical absorption spectrum and high microwave conductivity indicative of itinerant electrons. At raised temperatures the material is weakly paramagnetic (confirming the presence of unpaired electrons). But at low temperatures there is only residual diamagnetism, showing that the vast majority of electron spins are paired in the electride salts as the temperature approaches 0 K.

There now exists, therefore, the first example of a direct link between the spin-pairing interactions observed in liquid metal solutions and electron-pair trapping in these crystalline solids. The crystal structure of the electride $K^+(cryptand[2.2.2]) \cdot e^-$ shows the presence of large vacancies of complex shape, interconnected in two directions by zig-zag channels, but blocked in the third direction. Dye and co-workers also identify possible electron-pair trapping sites in the large cavities of the cryptate at the intersection of the various channels.

These significant results open up an attractive and versatile route to probing the nature of electron-pair localization versus itinerancy in condensed phases. At present, none of these solids has been observed to become superconducting at low temperature. In the light of the present excitement about high-temperature superconductivity, however, I would like to point out an important historic link between metal-ammonia solutions and the phenomenon of superconductivity at liquid-nitrogen temperatures.

The first proposal that mobile electron pairs were the cause of high-temperature superconductivity was made in 1946, not

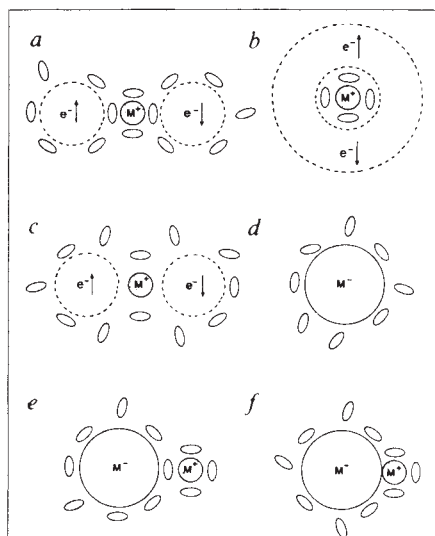


Fig. 1 The possible structures for diamagnetic (electron-spin-paired) species in solutions of alkali metals in non-aqueous solvent. The ellipsoids represent solvent molecules, the dotted lines represent the approximate extent (delocalization) of the electron wavefunction and the arrows represent the spin alignment. a, Loose triple ion; b, spherically symmetrical dianion of the solvated cation; c, contact triple ion; d, solvated alkali anion; e, alkali anion in a solvent-separated ion pair; f, alkali anion in a contact ion pair.

by a theoretical physicist, but by an experimental chemist, R. A. Ogg Jr of Stanford University⁷. This controversial proposal, arising out of Ogg's studies on magnetism and superconductivity (at 77 K) in quenched sodium-ammonia solutions⁷ was advanced more than a decade before the BCS theory⁸ of superconductivity. Ogg's idea was ingenious; because of their zero angular momentum (spin pairing), electron pairs in sodium-ammonia solutions must obey Bose-Einstein statistics. He proposed that rapid quenching of these metal solutions gives rise to high concentrations of trapped electron pairs and furthermore, that "such a state must

display the phenomenon of superconductivity". The new study¹ by Dye and co-workers will reopen the search for high-temperature superconductivity in materials derived from the liquid state. □

1. Huang, R.H., Faber, R.K., Moeggenborg, K.J., Ward, D.L. & Dye, J.L. *Nature* **331**, 559–601 (1988).
2. Edwards, P.P. *Adv. inorg. Chem. Radiochem.* **25**, 135 (1982).
3. Kraus, C.A. *J. Am. chem. Soc.* **30**, 1323 (1908).
4. Taylor, N.W. & Lewis, G.N. *Proc. natn. Acad. Sci. U.S.A.* **11**, 456 (1925).
5. Dye, J.L. *Prog. inorg. Chem.* **32**, 327–441 (1984).
6. Dye, J.L. & Ellaboudy, A.S. *Chem. Br.* 210–215 (1984).
7. Ogg, R.A. Jr *Phys. Rev.* **69**, 243–245 (1946).
8. Bardeen, J. *et al. Phys. Rev.* **106**, 162–164 (1957).

Peter P. Edwards is in the Department of Chemistry and the University Research Centre in Superconductivity, Cambridge University, Cambridge CB3 0HE, UK.

Palaeoanthropology

The dates of Eden

Chris Stringer

FOR the past half century, the relationship between Neanderthal and early modern human populations in south-west Asia has been a subject of considerable discussion and disagreement¹. On page 614 of this issue², Helene Valladas and colleagues report thermoluminescence dates of 92,000 years for the Mousterian levels of Qafzeh cave (Israel), which suggest that early modern humans were living there twice as long ago as is generally believed. These dates indicate that some modern *Homo sapiens* preceded Neanderthals in the area, standing the conventional evolutionary sequence on its head.

Publication of the first detailed descriptions of fossil human remains from the Mount Carmel caves in Israel³ began a long-standing controversy about evolutionary events in south-west Asia. The late Pleistocene caves of Tabūn and Skhūl contained hominid specimens which appeared to be related to both European Neanderthal (late 'archaic *H. sapiens*') and Cro-Magnon (early 'modern' *H. sapiens*) samples. Because of morphological variation and an uncertain relative chronology, it was initially assumed that these specimens represented a single highly variable population. Interpretations of their supposedly mixed or transitional features centred on ideas that they were either in the throes of evolutionary change or were the result of hybridization between distinct Neanderthal and modern populations. Subsequent work on associated mammalian faunas suggested that the Tabūn fossils were in fact older than those from Skhūl^{4,5}, and further study of the hominids indicated that they would be better regarded as distinct Neanderthal (Tabūn) and early modern *H. sapiens* (Skhūl) samples^{5,6}.

Discoveries from caves such as Shanidar (Iraq), Amud (Israel) and Kebara (Israel) have added considerably

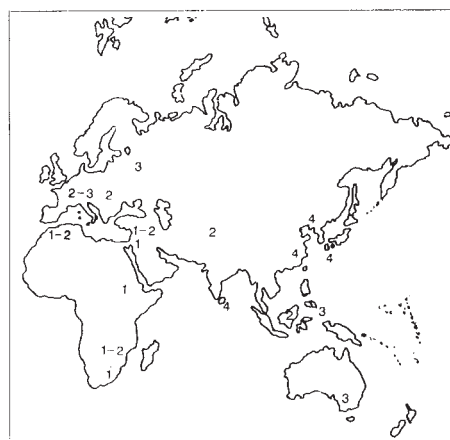
to the local Neanderthal sample, whereas Qafzeh (Israel) (see figure) has provided a further collection of early modern *H. sapiens* skeletons^{1,7,8}. Various workers, following the European hominid succession, have assumed that there was also a sequence from Neanderthal to modern *H. sapiens* populations in south-west Asia. This view has been supported by archaeological relative dating based on characteristics of the associated Mousterian stone tools⁹. The few reliable radiocarbon dates available suggested that some Neanderthals in the area were probably older than 50,000 years, and this view has been reinforced by thermoluminescence dates on Mousterian burnt flints from Kebara, dating the Neanderthal skeleton from there to about 60,000 years ago¹⁰.

Following the assumption that modern *H. sapiens* must have succeeded Neanderthals in the area, most people have favoured an estimated age for the Skhūl and Qafzeh hominids of about 40,000 years, but a few have argued¹¹ that the Qafzeh sample, at least, was much older. Biostratigraphy based on microfauna suggested that the Mousterian levels of Qafzeh were earlier than the main

Mousterian levels of Tabūn and Kebara, and it was proposed that the Qafzeh hominids might actually predate the Neanderthals. This idea found little favour, and other workers either disregarded the biostratigraphy¹ or, like me, preferred to sit on the fence and wait for further evidence. This evidence has now come in, and the minority view has been vindicated in a most startling fashion by thermoluminescence age determinations on burnt Mousterian flints from Qafzeh which average $92,000 \pm 5,000$ years before present. If accurate, these dates double the accepted age of the associated Qafzeh hominids.

The palaeoanthropological implications of such an age are enormous. First, it appears to establish that some anatomically modern humans (admittedly with primitive features) did indeed precede Neanderthals in south-west Asia. Evolutionary models centred on a direct ancestor-descendant relationship between Neanderthals and modern *H. sapiens* must surely now be discarded, along with associated schemes designed to explain such a transition. Second, a minimum chronological overlap between modern *H. sapiens* and Neanderthals of about 60,000 years in Eurasia, coupled with only highly disputed evidence of gene flow between them, reinforce calls for a return to a greater taxonomic separation than that of mere subspecies of *H. sapiens*^{12,13}.

Another outcome of the new Qafzeh dates is that modern *H. sapiens* fossils are now known from Israel as early as they are known from southern Africa, and this may support claims for an early north African presence of modern humans at sites such as Dar-es-Soltane (Morocco) and Omo-Kibish (Ethiopia). Both palaeoanthropologist and geneticist proponents of the "Out of Africa" model, which proposes a late Pleistocene African origin for modern *H. sapiens*, have tended to favour a sub-Saharan home for the species^{14,15}, but some rethinking may be necessary. An African origin for modern humans still seems most probable, but the time of origin must now be pushed back well beyond the age of the Qafzeh hominids. The possibility



Some important sites with early modern *H. sapiens* fossils. 1 (probably >50,000 years old), Klasies River mouth caves (South Africa); Omo-Kibish (Ethiopia); Qafzeh (Israel). 1–2 (possibly >50,000 years old), Border cave (South Africa); Dares-Soltane (Morocco); Skhūl (Israel). 2 (probably >30,000 years old), Mladeč (Czechoslovakia); Darra-i-Kur (Afghanistan); Niah cave (Sarawak, Malaysia). 2–3 (possibly >30,000 years old), Cro-Magnon (France). 3 (probably >20,000 years old), Sungir (Soviet Union); Tabon (Philippines); Mungo/ Willandra lakes (Australia). 4 (probably >10,000 years old), Batadomba lena (Sri Lanka); Liujiang (China); Zhoukoudian Upper Cave (China); Minatogawa (Japan).

exists that by 90,000 years ago the original *H. sapiens* stock had already split into (at least) a southern African component (represented at the South African caves of Klasies River mouth and, perhaps, Border cave) and a northern group (represented at Omo-Kibish, and extending to Qafzeh). Possible ancestral populations were present in both areas (sampled, for example, at the sites of Florisbad in South Africa and Jebel Irhoud in Morocco) as well as in between (for example, the fossils from Ngoloba in Tanzania and Omo-Kibish 2 in Ethiopia). Claims for an origin of modern *H. sapiens* in south-west Asia itself¹⁶, based on the morphology of the fragmentary Zuttiyeh skull, must now be taken into consideration as well.

Whichever area is favoured as the original "Eden" for modern humans, there are still other problems posed by the early Qafzeh dates. If modern *H. sapiens* populations were present in Israel 90,000 years ago, why did they take another 50,000 years to register their presence in Europe and the Far East? Were environments to the north so unsuitable or Neanderthals so well established that they prevented an early modern human radiation until much later? Or are relatives of the Qafzeh people waiting to be discovered elsewhere in Eurasia, part of a hitherto unsuspected early dispersal of modern *H. sapiens*? Whatever the answer, it now seems less likely that the Qafzeh sample bears the special relationship to the much younger Cro-Magnons of Europe implied by the name 'Proto-Cro-Magnons'. Finally, it is not possible to discuss here the considerable archaeological implications of the new Qafzeh dates, but it is more evident than ever that the behavioural changes which heralded the arrival of the peoples of the Upper Palaeolithic cannot be invoked to help explain the much more ancient origin of modern humans. □

1. Trinkaus, E. in *The Origins of Modern Humans* (eds Smith, F.H. & Spencer, F.) 251 (Liss, New York, 1984).
2. Valladas, H. *et al. Nature* **331**, 614–616 (1988).
3. McCown, T.D. & Keith, A. *The Stone Age of Mount Carmel II: The Fossil Human Remains from the Levallois-Mousterian* (Clarendon, Oxford, 1939).
4. Higgs, E.S. *Proc. prehist. Soc.* **27**, 144–154 (1961).
5. Howell, F.C. in *Hundert Jahre Neanderthal* (ed. von Koenigswald, G.H.R.) 185–198 (Kemink en Zoon, Utrecht, 1958).
6. Brothwell, D.R. *Proc. prehist. Soc.* **27**, 155–159 (1961).
7. Bar-Yosef, O. *et al. Curr. Anthropol.* **27**, 63–64 (1986).
8. Vandermeersch, B. *Les Hommes Fossiles de Qafzeh (Israel)* (CNRS, Paris, 1981).
9. Jelinek, A.J. in *The Transition from Lower to Middle Paleolithic and the Origin of Modern Man* (ed. Ronen, A.) 57–101 (BAR, Oxford, 1982).
10. Valladas, H. *et al. Nature* **300**, 159–160 (1987).
11. Bar Yosef, O. & Vandermeersch, B. in *Préhistoire du Levant* (eds Sanlaville, P. & Cauvin, J.) 281–285 (CNRS, Paris, 1981).
12. Stringer, C.B. in *Ancestors: the Hard Evidence* (ed. Delson, E.) 289–295 (Liss, New York, 1985).
13. Tattersall, I. *J. hum. Evol.* **15**, 165–175 (1986).
14. Gowlett, J. *Antiquity* **61**, 210–219 (1987).
15. Lewin, R. *Science* **237**, 1292–1295 (1987).
16. Vandermeersch, B. in *Les Processus de l'Hominisation* (ed. Ferembach, D.) 97–100 (CNRS, Paris, 1981).

Chris Stringer is in the Department of Palaeontology, British Museum (Natural History), Cromwell Road, London SW7 5BD, UK.

Supernova 1987A

The case of the speckled partner

E.S. Phinney

THE most striking of the unexpected features of the supernova 1987A, which occurred on 23 February last year in the Large Magellanic Cloud (see right), was the apparent creation of a bright red companion — the 'speck'. This speck, reported^{1,2} by two groups, appeared a month after the explosion, at a projected distance of 16 light days (4×10^{16} cm) from the supernova. It was visible for only a few weeks, and has not been detected in the many subsequent observations by its discoverers. In red light, it was about one-tenth as bright as the supernova, and 100 times brighter than the brightest stars in the vicinity before the explosion. Such bright objects appear in the heavens only rarely, and the probability that the speck was unrelated to the supernova is about 10^{-14} . The signal from the supernova which caused the speck must have travelled at more than half the speed of light. Obvious possible signals are light itself, and particles moving at relativistic speeds. The source of the speck must lie somewhere along a line segment beginning at the instruments which observed it and extending back to a point 10 light days behind the supernova (Fig. 1). What could this mysterious object be, and how could the supernova have caused it?

The first possibility to consider is that the speck was merely a mirage created in our distorting window on the Universe, the Earth's atmosphere. Hot, low-density cells of air rise buoyantly as cold, dense cells fall about them. Wind shear and convection twist these cells into turbulent eddies on scales of millimetres to kilometres. The index of refraction is lower in the hotter, more rarefied eddies and, as the originally planar wavefront from a distant star passes through the atmosphere, it runs faster in low-density cells and lags in dense ones. When it reaches our telescopes, the wavefront is corrugated, with points on the front separated by more than a distance $r_0 \approx 10$ cm shifted by several wavelengths (λ) of light. The apparent direction to the star will therefore differ from the true one by $\theta_0 \approx \lambda/r_0 \approx 5 \times 10^{-6}$ rad (1 arc s).

As the density fluctuations are swept along by the winds (velocity, $v \approx 5$ m s⁻¹), the corrugation of the wavefront will change on a timescale $\tau_0 \approx r_0/v \approx 0.02$ s — the familiar twinkling of the stars. In most astronomical observations, light is collected for much longer than τ_0 , so that incoming rays from every source are smeared over an angle $\approx \theta_0$. The smeared images of two sources separated by an angle less than $\theta_0 \approx 1$ arc s will be

blended, no matter how large the telescope. (The diffraction limit for the 4-m telescopes that discovered the speck is 0.03 arc s.)

So how can one hope to find objects like the speck just 0.06 arc s from the supernova? The clever method (introduced to astronomy by Labeyrie³) is to make many observations, each much shorter than the time τ_0 . To see how this allows us to reach the diffraction limit despite the atmosphere, consider a telescope covered by an opaque screen, with two holes smaller than r_0 and separated by a distance $a \gg r_0$. The telescope makes an image (which would otherwise appear at infinity) of the diffraction fringes produced when light waves leaving the two holes interfere. Because of the inhomogeneity of the Earth's atmosphere, the waves entering the two holes will typically be out of phase by a few wavelengths, and that phase delay will change on a timescale $\tau_0 \approx 0.02$ s, shifting the position of the diffraction fringes. In a long exposure no fringes will be visible, but for times much less than τ_0 , the phase difference between waves entering the two holes will be constant and fringes will be visible.

If light from two incoherent sources separated by an angle θ enters the holes, each source will produce its own fringes, and (in sufficiently short exposures) two superposed systems displaced relative to one another by $a\theta/\lambda$ fringes will be observed. Because of imperfections in the

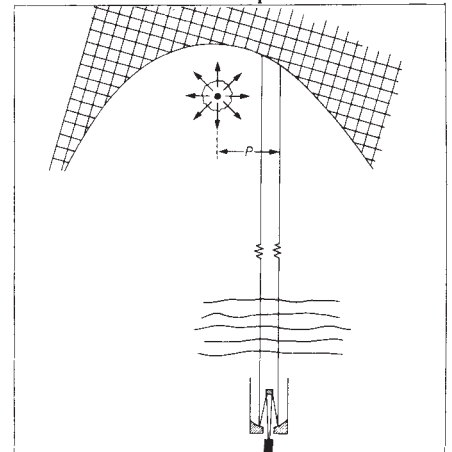


Fig. 1 The companion of SN1987A, detected by speckle interferometry, must lie somewhere along the double line extending from the instrumentation behind the telescope to a point 10 light days behind the supernova. $P = 16$ light days. Wavy horizontal lines indicate the Earth's inhomogeneous and refracting atmosphere. No signal from the supernova can travel to the hatched region and excite emission which could be received at Earth within 30 days of the first detection of the explosion.

Supernova 1987A: one year old

ON 23 February 1987, Supernova 1987A exploded in the Large Magellanic Cloud (LMC), about 55 kpc away. A brief flash of neutrinos, signifying the formation of a neutron star, has been followed during last year by emission from the expanding remnant in every part of the electromagnetic spectrum. The only radiation not seen was the burst of gravitational waves from the explosion; none of the detectors which might have seen it was operating.

Many of the observations and their theoretical implications could uncharitably be described as routine. The number and energy of the neutrinos conformed with models of the explosive energy release. When, after some confusion, the progenitor star had been identified as the blue supergiant Sk-69 202, the luminosity and radio emission around peak magnitude fitted well with complex hydrodynamical calculations of the core collapse and outburst. The subsequent exponential decline of the light curve was ample demonstration that decay of ^{56}Co was the energy source.

But minute scrutiny of astrophysical theories is rare, and the wealth of new and

detailed data from SN1987A has allowed already successful models to be refined by orders of magnitude. Twenty neutrinos on magnetic tape are worth any number of particle fluxes calculated by computer. Monitoring the light curve to unprecedented accuracy over many magnitudes allows exact deduction of the quantity of ^{56}Co synthesized, and of the presence or absence of secondary energy sources.

Now, as the expanding remnant thins and the core becomes visible, entirely new observations become possible. The γ -ray and infrared lines recently detected would not have been seen from a more distant supernova, and will allow astronomers, during the coming year, to watch the cooling remnant turn to gas and dust.

By following the day-to-day evolution of this supernova nebula, we will learn not only about the fate of one star, but about local conditions in the LMC, about interstellar matter between ourselves and the LMC, and with greater certainty about supernovae in distant galaxies, of which two have already been spotted this year.

David Lindley

telescope optics, the shape of the single-source fringe pattern is not known *a priori*, but must be determined by observing a fiducial star known to be a point source. A difference between the fringe pattern of the source of interest and that of the fiducial star is then evidence for structure in the source. Of course there will always be statistical differences, and if the fiducial star and the source do not follow sufficiently similar paths through the atmosphere and telescope optics, there could be systematic differences as well. These might be misconstrued as evidence for a spurious faint second source (which would produce only a small modulation of the fringes of the brighter source).

The observational technique (speckle interferometry) which discovered the speck is more sophisticated than the simple two-hole system described above (the telescope is not masked at all), but the principles, and limitations, are the same. Although the technique is reliably used to measure the separations of close double stars, the many possibilities for systematic error are highlighted by occasional spurious discoveries (like the 'companion' to the star VB8, which inspired a whole conference before it was shown not to exist).

There is yet another way that SN1987A could have created a mirage: its light could have been deflected by a compact object between it and the Earth¹. The mass required ranges from about 10 solar masses ($M_{\odot}$) if in our Galaxy, to about $10^5 M_{\odot}$ if near the supernova's galaxy. The numbers of such hypothetical compact objects are, however, constrained by the known total

masses of the galaxies, and the probability of producing a second image of the supernova in this way is less than 10^{-6} . The reported colour difference between speck and supernova is a further problem for this model. If the speck was real, it must have shone by reflecting the light of SN1987A or by emitting a line spectrum (from rarefied gas), a nonthermal continuous spectrum (from relativistic particles) or a thermal continuous spectrum (from dense gas or a hot solid). People are circulating preprints invoking all these mechanisms, but only the thermal spectrum is consistent with the observations as reported, except in implausible situations.

Our view of interstellar space is obscured by small grains of condensed refractory elements. A large cloud of these (perhaps the disk of an inchoate companion star) could survive heating by the light pulse of SN1987A and reflect some of its light. Difficulties arise, however, because dynamical arguments and the observed asymmetry make it unlikely that the cloud could cover more than one-tenth of the sky of SN1987A. An uncomfortably high bulk albedo is thus required. The problem is even worse if the reported colours of the speck¹ are accepted. Grains scatter more blue light than red in the forward direction, so it is hard to believe that a cloud in front of the supernova reflects three times more red light than blue, as indicated by the observations. Light scattered through the cloud can be red enough, but would be too faint. If the cloud were behind SN1987A, blue light would penetrate deeper into the cloud and be absorbed. A disk or grains with conventional proper-

ties thus reflects about 30 per cent of the incident red light, but only about 22 per cent of the blue. But were the cloud behind the supernova, scattered light arriving at Earth 30 days after the explosion would have to have been emitted by SN1987A less than 15 days after the explosion, when it was much fainter and bluer. The cloud would have to reflect almost all of the incident red light, but only 2–17 per cent as much of the blue light, and no plausible fiddling of grain parameters can produce these bulk albedos.

Line emission⁵ is excluded by high-resolution spectra⁶. Speckle observations show that in the 100-Å band around the hydrogen H α emission the speck was one-tenth of the brightness of SN1987A. If the speck's one-tenth of the light in this 100-Å band were concentrated in a line less than 40-Å wide, it would easily be recognized in the combined spectrum; careful fitting could identify even broader lines. In its first day, SN1987A probably emitted a pulse of about 4×10^{46} erg of ionizing radiation (refs 7, 8 and D. Arnett, preprint) which could ionize a dust-free gas cloud. Models of this sort, invoking the recombination of ions with electrons, are hard pressed to produce the observed H α luminosities⁹. The mass of ionized gas required is of the order of $M_{\odot}$, and the expected linewidth is only about 0.4 Å (caused by Doppler broadening by speeds of 20 km s⁻¹). Even with an ionizing pulse orders of magnitude larger than expected (so that the ionized cloud becomes thick to Compton scattering⁹), or using shock waves to do the ionization, it does not seem possible to get the line wider than

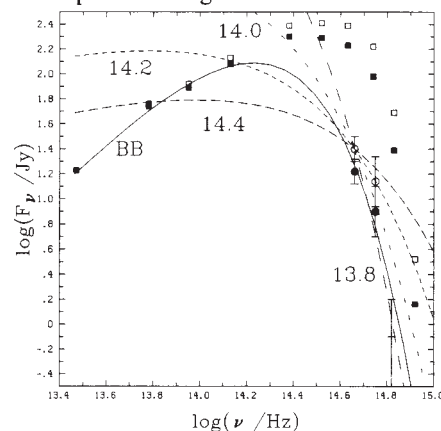


Fig. 2 Spectrum of SN1987A plus speck (squares) and speck alone (circles) 30–37 days after the explosion. Filled points are as observed¹⁶; open points have been de-reddened assuming the relative extinction of blue and visible light (E_{B-V}) is 0.2, and the extinction curve is that of the Large Magellanic Cloud¹⁴. Dashed curves, single-particle synchrotron emission curves with critical frequencies as labelled. Note that no synchrotron spectrum is simultaneously compatible with the total flux measurements in both the infrared and ultra-violet. The curve labelled BB, emission from a 3,000 K black body, is marginally consistent with the total flux in the infrared.

500 km s⁻¹, a factor of four short of that needed to prevent the line from appearing in the combined spectrum.

Models of the speck invoking synchrotron radiation^{10,11} or bremsstrahlung (S.A. Colgate *et al.*, preprint) are inconsistent with the accurately known spectrum of the combined light from SN1987A and the speck (Fig. 2). Any synchrotron or bremsstrahlung emission which fits the reported colours of the speck, produces an infrared flux at 5 μ m that is at least 30 times higher than that observed. Even if one ignores the reported detections of the speck at wavelengths other than 6,563 Å and allows arbitrary amounts of dust local to the emitting region, such models produce either more far-infrared emission, or more ultraviolet light than observed from the supernova. Synchrotron and bremsstrahlung models would be viable only if the speck were more than four times fainter than reported. The lack of radio emission and the low polarization¹² pose additional problems for these models.

The only emission mechanism without severe difficulties is thermal emission from a photosphere or hot dust. Temperatures of 3,000–4,000 K are required; the photosphere would have a radius of about 10¹⁵ cm. Surprisingly, exotic physics seems to be needed for this most prosaic of emission mechanisms to work. Among the models proposed are marble-sized dust lumps melting in the radiation field of a ultraviolet flash thousands of times brighter than predicted by models of the supernova light curve (J.J. Quenby, preprint), a protostellar gas cloud heated by a 10⁵¹ erg relativistic precursor or by cosmic rays streaming ahead of the supernova shock, or a piece of neutron star ejected at relativistic speed by the supernova explosion (ref. 13 and E.D. Carlson, S.L. Glashow and U. Sarid, preprint) yet somehow prevented from expanding faster than 3,000 km s⁻¹.

None of these models seems immediately compelling. Without further observations of the speck, or detection of the unusual X-ray emission that would occur as the supernova ejecta struck the protostellar gas cloud predicted by some of the models, the companion of SN1987A may remain one of astronomy's curiosities. □

1. Nisenson, P., Papaliolios, C., Karovska, M. & Noyes, R. *Astrophys. J.* **320**, L15–L18 (1987).
2. Meikle, W.P.S., Matcher, S.J. & Morgan, B.L. *Nature* **329**, 608–611 (1987).
3. Labeyrie, A. *Astr. Astrophys.* **6**, 85–87 (1970).
4. Barrow, J.D. & Subramanian, K. *Nature* **327**, 375 (1987).
5. Hillebrandt, W. *et al. Astr. Astrophys.* **186**, L9–L10 (1987).
6. Menzies, J.W. *et al. Mon. Not. R. astr. Soc.* **227**, 39P (1987).
7. Falk, S.W. *Astrophys. J.* **225**, L133–L136 (1987).
8. Woosley, S.E. *et al. Astrophys. J.* **318**, 664–673 (1987).
9. Felten, J.E. *et al. Astrophys. J.* (in the press).
10. Rees, M.J. *Nature* **328**, 207 (1987).
11. Blandford, R.D. *Proc. IAU Colloq.* **101** (in the press).
12. IAU Circ. 4358, 4361.
13. Piran, T. & Nakamura, T. *Nature* **330**, 28 (1987).
14. Clayton, G.C. & Martin, P.G. *Astrophys. J.* **288**, 558 (1985).

E.S. Phinney is in the Division of Physics, Mathematics and Astronomy, California Institute of Technology, Pasadena, California 91125, USA.

Cell motility

Myosin filaments on the move

J.V. Small

THOSE who were impressed by the elegantly simple demonstration^{1,2} that single myosin heads bound to a substrate can alone translocate actin filaments may doubt that the rest of the molecule (Fig. 1b) is necessary for motility. But investigations of intact non-muscle and smooth muscle myosins have consistently shown that filament formation is closely coupled to the expression of the actin-activated magnesium ATPase activity of the myosin

assembly, as does tail-end phosphorylation³. Although the overall pattern is complicated by the presence or absence of phosphorylation sites, the tail-end of myosins from vertebrate non-muscle cells^{5,9}, from smooth muscle¹⁰ and from skeletal muscle¹¹ have similarly been implicated in filament formation.

Even more intriguing are the effects of myosin heads on filament assembly and disassembly. In those systems where

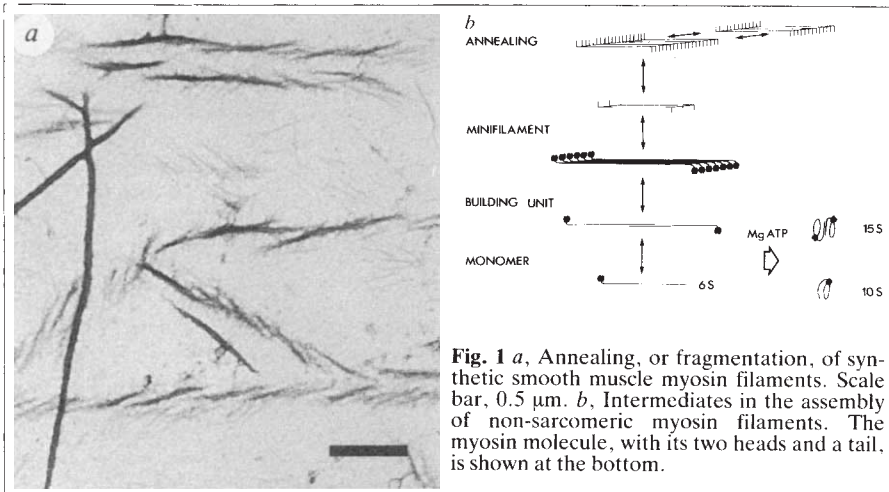


Fig. 1 *a*, Annealing, or fragmentation, of synthetic smooth muscle myosin filaments. Scale bar, 0.5 μ m. *b*, Intermediates in the assembly of non-sarcomeric myosin filaments. The myosin molecule, with its two heads and a tail, is shown at the bottom.

molecule, which is essential for movement. Interesting and important aspects of the assembly of these myosins are now illuminated by several studies^{4–7}. The new results not only document the dynamic molecular-exchange mechanisms between the polymeric and monomeric pools of the molecules, but reinforce earlier conclusions about the structural uniqueness of non-sarcomeric myosin filaments.

We are now witnessing the coming of age of studies on myosin assembly, with promises of as many interesting surprises as have been yielded by its contractile partner actin over the past decade. Moreover, with the increasing consensus that thick filaments in other systems may not after all be constructed like those in skeletal muscle⁸, the way lies open for some new ideas about non-muscle cell motility and smooth muscle contraction.

Two sites on the myosin molecule (see Fig. 2) appear to play key roles in the assembly of molecules into filaments; one is close to the far end of the tail and the other is in the 'neck' region of the head. Attention was first drawn to the tail end by the finding that assembly of myosin into thick filaments in amoebae is controlled by reversible phosphorylation of several amino-terminal serine residues. Enzymatic cleavage of the tail piece or antibodies against this region inhibit filament

calcium-dependent regulation of contraction or of locomotion operates via phosphorylation of regulatory light chains on the heads, phosphorylation shifts the monomer–polymer equilibrium, *in vitro*, in favour of assembly. Under appropriate conditions and in the presence of micromolar concentrations of MgATP, light-chain dephosphorylation promotes filament disassembly and, remarkably, folding of the myosin tail into three, first shown for vertebrate smooth muscle myosin¹² and subsequently for myosins from vertebrate non-muscle cells¹³. The headless myosin rod alone cannot perform the same acrobatics; rod filaments are stable even in millimolar ATP. Two other myosins, from *Physarum*⁹ and now from an invertebrate smooth muscle¹⁴ also fold, but here phosphorylation of the rod, presumably close to the amino terminus of the molecule, is involved.

The recent studies of Cross *et al.*, of Trybus and Lowey and of Kendrick-Jones and co-workers reveal new aspects of the filament assembly process. Trybus and Lowey¹ define ATP-free conditions under which smooth muscle myosin, whether or not it is phosphorylated, forms homogeneous bipolar minifilaments comparable to those earlier produced from skeletal muscle myosin by Reisler. The extraordinary clarity of the shadowed

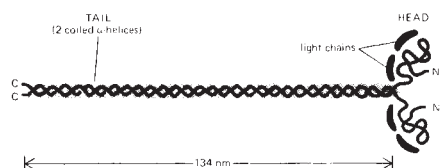


Fig. 2 Schematic drawing of a myosin molecule showing the two heavy chains (tail) and four light chains (head). The ATPase is contained in the heads.

images of filaments in the electron microscope readily allows a head count of the number of myosin molecules, 12–14 in each filament, and indicates a longitudinal spacing between the heads of around 43 nm. Using the further observation that the smooth muscle myosin rod also forms minifilaments under the same conditions, Trybus and Lowey demonstrate, in mixing experiments, a rapid subunit exchange between filaments occurring within minutes. By the same mechanism, smooth muscle myosin minifilaments readily incorporate skeletal muscle myosin molecules. Using different conditions, Kendrick-Jones and co-workers⁶ independently demonstrate critical concentrations of monomer for assembly into filaments for smooth muscle, thymus and brush border myosins. The dynamic nature of myosin monomer–polymer transitions is thus brought into the limelight. Surprisingly, these transitions have also been seen in skeletal muscle myosin¹⁵.

But what about the folded state described above? Cross and colleagues⁷ show that folding is concomitant with ATP hydrolysis and causes the trapping of the hydrolysis products in the myosin head. They demonstrate that the folding reaction blocks not only the ATPase but also the self-assembly of myosin molecules so that the folded molecules represent an inert storage form of myosin. This result readily explains the apparent increase in critical concentration for polymerization induced by MgATP⁶ — the critical concentration for polymer-competent myosin remains unchanged against an increased background of redundant folded monomers. It is not known how these myosins fold or what amount of folded myosin occurs in systems such as smooth muscle *in vivo*. But some interesting transition forms to the folded state are providing clues about filament construction (Fig. 1). Trybus and Lowey^{3,4} and Onishi and Wakabayashi¹² have identified a stable 15S folded dimer that apparently unfolds to form an anti-parallel bipolar unit with a tail overlap of 40–50 nm. Such a building unit, as suggested earlier¹⁰, would explain the characteristic absence of a central bare zone and the 'mixed' or 'side-polar' appearance of long filaments assembled from smooth and non-muscle myosins^{8,16,17}.

How does myosin head phosphorylation fit into the picture? It is generally

agreed that phosphorylated myosin filaments, apart from being enzymatically competent to interact with actin, are more stable. But the new results differ concerning whether all or only a proportion of phosphorylated molecules can confer resistance to solubilization by MgATP. Trybus and Lowey find⁴ that dephosphorylated smooth muscle myosin is readily dissociated from smooth muscle minifilament copolymers containing phosphorylated and dephosphorylated myosin, whereas Citi *et al.* observe⁵ stabilization in comparable non-muscle myosin copolymers containing as little as 30 per cent phosphorylated myosin. In the latter case, however, Citi *et al.* used longer filaments rather than minifilaments, which could be significant.

As is clear from the images of Trybus and Lowey³, the myosin heads in the minifilaments are all far removed from the overlapping tails and cannot directly affect tail–tail interactions. It is conspicuous, however, that the growth of smooth muscle minifilaments into larger filaments occurs first laterally, reminiscent of the side-to-side, slightly oblique stacking of *Acanthamoeba* myosin minifilaments¹⁸. This staggered annealing of minifilaments leads to the population of more crossbridge levels and to the appearance of the characteristic 14-nm myosin helical repeat^{8,16,18} that can, with subsequent growth, extend to the centre of the filament⁸. Once this occurs, in filaments longer than about 300 nm, head–tail interactions, modifiable by head phosphorylation, could take place. The filaments could then be stabilized by submaximal levels of myosin phosphorylation. A consequence of such a building mode is that filament growth need not be restricted to the addition of bipolar dimers but could just as well proceed by the annealing of more minifilaments or larger assemblies⁸ (Fig. 1a). □

1. Harada, Y., Noguchi, A., Kishino, A. & Yanagida, T. *Nature* **326**, 805–808 (1987).
2. Toyoshima, N.Y. *et al.* *Nature* **328**, 536–539 (1987).
3. Trybus, K.M. & Lowey, S.J. *Cell Biol.* **105**, 3007 (1987).
4. Trybus, K.M. & Lowey, S.J. *Cell Biol.* **105**, 3021 (1987).
5. Citi, S., Smith, R.C. & Kendrick-Jones, J. *J. molec. Biol.* **198**, 253–262 (1987).
6. Kendrick-Jones, J., Smith, R.C., Craig, R. & Citi, S. *J. molec. Biol.* **198**, 241–252 (1987).
7. Cross, R.A., Cross, K.E. & Sobieszek, A. *EMBO J.* **5**, 2637–2641 (1986).
8. Hinssen, H., D'Haese, J., Small, J.V. & Sobieszek, A. *J. ultrastruct. Res.* **64**, 282–302 (1978).
9. Kuznicki, J. *FEBS Lett.* **204**, 169–176 (1986).
10. Cross, R.A. & Vandekerckhove, J. *FEBS Lett.* **200**, 355–360 (1986).
11. Nyitrai, L. *et al.* *J. biol. Chem.* **258**, 13213–13220 (1983).
12. Onishi, H. & Wakabayashi, T. *J. Biochem., Tokyo* **92**, 871–879 (1982).
13. Craig, R., Smith, R. & Kendrick-Jones, J. *Nature* **302**, 436–439 (1983).
14. Castellani, L. & Cohen, C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4058–4062 (1987).
15. Saad, A.D., Pardee, J.D. & Fischman, D.A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9483–9487 (1986).
16. Sobieszek, A. *J. molec. Biol.* **70**, 741–744 (1972).
17. Craig, R. & Meggermann, J. *J. Cell Biol.* **75**, 990 (1977).
18. Pollard, T.D. *J. Cell Biol.* **95**, 816–825 (1982).

J.V. Small is at the Institute of Molecular Biology, Austrian Academy of Sciences, 5020 Salzburg, Austria.

Daedalus

Encapsulated breath

A FIRE on an aircraft kills more by smoke-inhalation than by flame. One defence is to provide the passengers with smoke-hoods that give a few minutes' protection. Daedalus now proposes a more fundamental approach. He points out that the absorption area of the stomach and intestines is about 20 per cent of that of the lungs. Because the lungs extract oxygen from air containing only 20 per cent of that gas, we should be able to 'breathe' quite effectively through the stomach and intestines if they were provided with pure oxygen.

To do its new job properly, the stomach would also have to dispose of respired carbon dioxide. In this connection Daedalus recalls that sodium peroxide reacts with carbon dioxide to release oxygen: not in perfect replacement-ratio, but enough to make it useful in air-purifiers for submarines. So DREADCO's chemists are exploring alkali-metal peroxide chemistry in search of a suitable composition for a 'stomach-breathing pill'. The best bet seems to be a balanced mixture of sodium superoxide and lithium oxide: it contains about 38 per cent of available oxygen, and releases it in exchange for the equivalent quantity of carbon dioxide.

These ferocious oxides cannot just be swallowed. They react energetically with water to form highly caustic alkalies. The DREADCO team is encapsulating its formulations into thousands of tiny silicone-rubber-coated granules, to make a sort of micro-sago or 'oxygenated caviar'. The silicone rubber safely retains the solid reagents and excludes water, but is permeable to oxygen and carbon dioxide. A gut charged with oxygenated caviar should smoothly take over the function of the lungs. A mere hundred grams of this powerful but insipid nostrum should enable the swallower to hold his breath effortlessly for at least an hour.

Thus the smoke-hood can be replaced by something only slightly less palatable than standard airline food. Gorged on the new emergency rations, passengers will be able to stroll to safety through the thickest smoke. Firemen will need no breathing apparatus. Even a victim already overcome by smoke might be revived by force-feeding with oxygenated caviar.

The new ability to stop breathing will also be welcomed by skin-divers, high-altitude mountaineers and beleaguered non-smokers. But stomach respiration will pose yet another chemical problem for sports authorities. Oxygenated both through lungs and stomach, athletes will break ever more records while denying that they are making use of unnatural stimulants.

David Jones

The molecular nature of the cystic fibrosis antigen

SIR—Cystic fibrosis is the most common autosomal recessive genetic disease among Caucasians. About 1 in 20 people, though healthy, is a carrier, and 1 in 2,000 newborns is affected by the disease. No reliable biochemical marker is available for the defective gene which has been localized on chromosome 7 (ref. 1). However, using isoelectric focusing, a serum factor (pI 8.4 and 8.5) has been identified² which is present at elevated levels in patients and obligate heterozygotes (parents). The gene for this factor, called the cystic fibrosis (CF) antigen, was assigned to chromosome 1 (ref. 3). Dorin *et al.*⁴ described the molecular cloning of a myeloid leukaemia protein which could be isolated with a monoclonal antibody for the CF-antigen. Recently we described the isolation and molecular cloning of a pair of proteins, MRP-8 and MRP-14, from peripheral blood leukocyte cultures⁵, which both map to chromosome 1 (K.H. Grzeschik, personal communication). The sequence of MRP-8 is identical to the sequence of Dorin *et al.*⁴ with the exception of one additional nucleotide, which alters the predicted last 15 amino acids, due to a shift in the reading frame.

With monospecific antisera against the recombinant proteins we are able to titrate MRP-8 and MRP-14 at levels of less than 1 ng ml⁻¹. Unexpectedly, on screening many sera and plasmas of carriers and non-carriers, we detected MRP-8 in only three cases (2–5 ng ml⁻¹), whereas we always found MRP-14. For cystic fibrosis patients, the plasma levels of MRP-14 were in the range 250–5,500 ng ml⁻¹ ($n = 16$), and for obligate heterozygotes in the range 120–380 ng ml⁻¹ ($n = 23$) with one sample at 75 ng ml⁻¹. For control individuals, values of 2–50 ng ml⁻¹ were found in 32 cases, and values of 55–135 ng ml⁻¹ in seven cases. These data suggest that MRP-14 is the CF-antigen. Because we isolated MRP-8 and MRP-14 as a complex⁵, Dorin *et al.* probably isolated the CF-antigen as a complex of MRP-8 and MRP-14 as well. If so, they obtained only MRP-8 sequence because, as shown by our own data, MRP-14 is N-terminally blocked. Another possibility, though remote, would be that a protein with the 15 C-terminal amino acids as reported⁴ represents a different allele of MRP-8 not recognized by our antiserum. In that case more than one CF-antigen would exist.

The increased level of CF-antigen(s) might not be unique to cystic fibrosis; in plasma of patients suffering from other diseases (for example polyarthritis) we also find levels well above the normal average. It will be intriguing to discover the common denominator.

In our preliminary study there is a

minor overlap between the MRP-14 plasma levels of obligate heterozygotes for the CF gene and other healthy volunteers. This overlap is probably caused by samples from unidentified carriers (1 in 20) and unrecognized disease cases. Nevertheless, it so far seems possible to distinguish most non-carriers from carriers of the CF gene, by a simple enzyme-linked immunosorbent assay (ELISA) test using our monospecific anti-MRP-14 serum. This will be important for relatives of CF patients.

JOSEF BRÜGGEN
LAJOS TRICSAY
NICO CERLETTI
KAREL ODINK

Department of Biotechnology,
Ciba-Geigy, K-681.546,
4002 Basel, Switzerland

MARKUS RUTISHAUSER
GEORG HOLLÄNDER

Children's Hospital,
4005 Basel, Switzerland

CLEMENS SORG

Department of Experimental Dermatology,
University of Münster,
44 Münster, FRG

1. Estivill, X. *et al.* *Nature* **326**, 840–845 (1987).
2. Wilson, G.B. *et al.* *Clin. chim. Acta* **49**, 79–91 (1973).
3. van Heyningen, V. *et al.* *Nature* **315**, 513–515 (1985).
4. Dorin, J.R. *et al.* *Nature* **326**, 614–617 (1987).
5. Odink, K. *et al.* *Nature* **330**, 80–82 (1987).

Eustatic sea-level change

SIR—Gaffin¹ has demonstrated a negative correlation, with a phase difference of 10 Myr, between geomagnetic reversal rate and recent data on long-term eustatic sea-level change². He continues, "it is reasonable to demand that the dominant mechanism controlling eustasy be plausibly linked to reversals too" and assumes that oceanic ridge volume is the controlling factor. His calculation of ocean basin volume takes other variables (such as continental hypsometry and water volume) to be fixed. With these constraints he calculates the seafloor creation rate required to cause the long-term eustatic sea-level change, and finds a positive correlation, again with a 10-Myr phase difference. Such a result is of great theoretical interest.

There are, however, other possible controls on sea level that need to be considered. For example: (1) change in the age distribution of oceanic crust has been shown to have an effect on ocean-basin volume^{3,4}. The development of 'Atlantic' type oceans (with no marginal subduction) at the expense of 'Pacific' type oceans (where the age of oceanic crust being subducted is essentially random) is not the same as a simple change in spreading rate (as suggested by Gaffin⁵), as the age distribution of non-subducted crust will change. The development of the Atlantic and Indian oceans over the past

150 Myr, the timespan considered by Gaffin¹, would give rise to a rise and subsequent fall in sea level of the order of tens of metres. A fall of roughly 100 m since the mid-Cretaceous has been suggested⁶.

(2) The similarity of modern normalized hypsometric curves⁶ and the apparent relationship between modern continental area and elevation^{7,8} suggests that former continental masses should have conformed to the same patterns. Using modern data on continental hypsometries⁹ it is possible to calculate the effect on global sea level of the split up of Pangaea⁸. This gives a predicted rise in sea level of some 130 m between about 140 and 50 Myr before present, and a small drop (of the order of 10 m) since then.

(3) Massive mid-plate volcanism from about 110 to 70 Myr before present in the Western Pacific has been used to suggest a sea-level rise at that time of a minimum of 40 m and probably nearer 100 m (ref. 10).

All these changes need to be removed from whatever eustatic sea-level curve is used before ascribing the remaining variation to change in ridge volume. The large differences in absolute values of sea-level change between some recently published curves^{2,11}, and the view that in any case we are unable to isolate eustatic changes from other sea-level change¹² suggest caution in applying Gaffin's correlations.

ANTONY R. WYATT

Department of Geology,
University College of Wales,
Aberystwyth, Dyfed SY23 3DB, UK

1. Gaffin, S. *Nature* **329**, 816–819 (1987).
2. Haq, B.U. *et al.* *Science* **235**, 1156–1166 (1987).
3. Berger, W.H. & Winterer, E.L. in *Pelagic sedimentation: on land and under the sea* (eds Hsu, K.J. & Jenkyns, H.C.) (Blackwell Scientific, Oxford, 1974).
4. Heller, P.L. & Angevine, C.L. *Earth planet. Sci. Lett.* **75**, 417–426 (1985).
5. Gaffin, S. *Am. J. Sci.* **287**, 596–611 (1987).
6. Harrison, C.G.A. *et al.* *Tectonics* **2**, 357–377 (1983).
7. Wyatt, A.R. *Nature* **311**, 370–372 (1984).
8. Wyatt, A.R. *J. geol. Soc. Lond.* **143**, 907–910 (1986).
9. Cogley, J.G. *Z. Geomorphol., Suppl.* **53**, (1985).
10. Schlanger, S.O. *et al.* *Earth planet. Sci. Lett.* **52**, 435 (1981).
11. Hallam, A. *Rev. Earth planet. Sci.* **12**, 205–243 (1984).
12. Burton, R. *et al.* *Earth Sci. Rev.* **24**, 237–277 (1987).

Any old iron?

SIR—Martin and Fitzwater (*Nature* **331**, 341–343; 1987) suggest that the growth of phytoplankton in ocean waters at high latitudes is restricted today by the limited amount of iron available, and that during recent ice ages the carbon dioxide content of the atmosphere was reduced because there was then more dust in the air. This suggests a possible way to alleviate the anthropogenic greenhouse effect, which is at present a cause for concern. By adding iron compounds to the oceans, a 'technological fix' to remove carbon dioxide from the air might be practicable.

JOHN GRIBBIN

Rose Cottage, Piddinghoe,
East Sussex BN9 9PT, UK

Neuroscience: a new era?

H.B. Barlow

Neural Darwinism: The Theory of Neuronal Group Selection. By Gerald M. Edelman. Basic Books: 1987. Pp.371. \$29.95.

THE power of Darwinian selection is much in the air these days, especially after the success of the clonal selection theory in immunology; it has, for instance, been applied to the travelling salesman problem by Brady (*Nature* **317**, 804–806; 1985) and to the learning of temporal sequences such as bird song by Dehaene and colleagues (*Proc. natn. Acad. Sci. USA* **84**, 2727–2731; 1987).

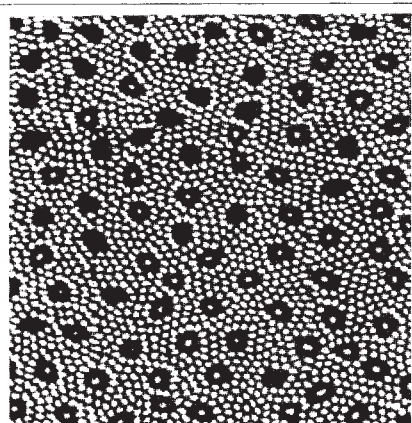
The basic idea of Gerald Edelman's book is stated in the title: neurons are organized into groups, and these groups are subject to a selection process which forms the anatomical and functional patterns of connection that are responsible for the astonishing performance of the brain. Assessment of the merit and originality of the idea depends upon the specifics of these proposals: what are neuronal groups, what are the selective processes applied to them, and where is the potential for exponential growth that is so essential for a Darwinian process to have the power it does? Finding answers to these questions is not easy, for although the distinguished author has obviously made desperate efforts to explain himself clearly, this is an exceedingly difficult book to read and understand.

One of the difficulties is stylistic. In discussing the localization of function in the brain, Edelman says (pp.141–142): "The position taken here (and the only one consistent with neuronal group selection theory) is that the real basis for overall functional responses is the dynamic interaction of specific individual components arranged in repertoires of neuronal groups or populations within different mapped reentrant structures rather than the fixed assignment of function to anatomically distinct regions". That sentence is hard to understand, and so are many others in the book.

A second difficulty lies in the lack of clarity and completeness in the definitions. For example, it is crucial to understand exactly what Edelman means by the term "neuronal group", which figures so prominently in his thinking. Perhaps the simplest question to ask is: are the groups overlapping, or non-overlapping? If a given neuron can belong to more than one group the number of possible groups is virtually unlimited and it might be possible to devise a rule for group reproduction that would give the potential for exponential growth. Both Darwin and Wallace appreciated that it was the

potential Malthusian population explosion that gave natural selection its immense power, and the glimpse of neuronal groups with these properties makes one say to oneself "Aha, this idea might work". If, on the other hand, the groups are non-overlapping, then reproduction and selection can merely shuffle neurons from one group to another, and it is hard to believe that this contains any such exciting potential.

Eagerly turning to the book, one finds the primary definition of a neuronal group on pp.46–47, as follows: "... a collection of cells of similar or variant types, ranging in number from hundreds to thousands, that are closely connected in their intrinsic circuitry and whose mutual dynamic interaction may be further enhanced by



Seeing spots — a section through the peripheral retina of a monkey, passing through the layers of rods (small white dots) and cones (black regions with white dots in centre). The picture is taken from *Eye, Brain and Vision* by David H. Hubel, recently published in Britain by W. H. Freeman.

increases in synaptic efficacy". This does not resolve the issue, and it is not for another 150 pages or so that it becomes reasonably clear that, alas, the groups do not overlap. Such an elementary point upon which so much hinges should have been resolved with the first definition.

It is disappointing that groups appear to be non-overlapping, but it is hard to see the merit of the overall concept for other reasons as well. Edelman says (pp.164–165): "According to the model, a neuronal group in the cerebral cortex is functionally defined as an ensemble of cohesively interconnected cells, all of which express the same receptive field". In the visual cortex there are no reports of two or more different cells having identical receptive fields. But apart from the lack of experi-

mental evidence for such groups, what theoretical or practical advantage could be conferred by having a set of cells with identical function?

The selection process is the next hurdle. This depends upon synaptic modification according to a Hebbian rule, as in so many current neural network models (for example, Longuet-Higgins *et al. Q. Rev. Biophys.* **3**, 223–244; 1970 and *Parallel Distributed Processing*, by McClelland, Rumelhart and the PDP research group, MIT Press; 1986). What is curious, however, is that in Edelman's version the result of positive selection is to incorporate more and more cells into the same group. What is the benefit of this? It is not like the increase in population of a species, which implies that they have successfully competed for a larger fraction of the available primary resources so that the genes responsible for the success have multiplied. For more and more cells in one group to respond to the same stimulus confers no particular advantage either to the cells, or to the group, and nothing corresponding to the genes has been proposed. Such a process would probably wreck the function of the whole brain, for if the cells in a group respond to a larger fraction of the sensory stimuli being received, this implies that they are becoming less selective: the brain would simply become less discriminating, more muddled. Furthermore this is not what happens in the developing visual cortex (Blakemore & Vital-Durand, *J. Physiol.* **345**, 40P; 1983 and Derrington, *Exp. Brain Res.* **55**, 431–437; 1984) where experience makes the cells become more, not less, selective. Thus Edelman's type of selection does not seem to happen, and it is hard to see how it would confer any advantage if it did.

The book contains an interesting section on cell adhesion molecules and their effects in development, but this is not a complete review of all the work in the field. There are also accounts of two computer simulations of the processes Edelman thinks are at work in the brain, and these have the merit of trying to give greater precision to the ideas that his words leave unclear.

This is not a book for the faint-hearted, because of its difficulty and obscurity, nor is it for the untutored, because it gives such an incomplete and one-sided account of brain development and function. But epoch-making books are often panned unfairly by bigoted reviewers, so readers are strongly urged to study "this magisterial work" (as the publisher's blurb calls it) and decide for themselves whether it ushers in a new era in neuroscience, or whether it's just a hopeless muddle. □

H.B. Barlow is Royal Society Research Professor in the Physiological Laboratory, University of Cambridge, Downing Street, Cambridge CB2 3EG, UK.

Digging against the clock

Paul G. Bahn

Recent Archaeological Discoveries in Japan. Edited by Tsuboi Kiyotari. Translated by Gina L. Barnes. *Unesco/Centre for East Asian Cultural Studies*: 1987. Pp.127. Pbk FF.100, £15, \$23.

THIS, the latest addition to an attractive series that has been appearing since 1979, is as informative and interesting as its predecessors, which covered recent archaeological discoveries in China, Korea and India. The pace of production and publication of the volumes is slow; but as a survey of work in Japan up to 1984, the book constitutes a readable and well-illustrated account for those without easy access to or comprehension of the Japanese literature.

One's faith in its accuracy is shaken by the very first sentence of the Introduction (which states that "the Japanese archipelago lies west of the Asian landmass"), but happily this is a minor slip in an otherwise excellent text. Until recently, Japan was much neglected in archaeology courses taught in other parts of the world, in part through the unavailability of translated material. Yet its location in relation to a continent is not dissimilar to that of Britain on the other side of Eurasia (both, for example, became islands through the postglacial rise in sea level), a point which should perhaps have stimulated more comparative studies. In addition, because Japan extends from the subarctic to the subtropical zone, past human adaptations to very different environments can be studied within a fairly small distance.

The book contains chapters on all of the main periods recognized in Japanese archaeology, each written by a specialist. The treatment is somewhat uneven, a situation dictated partly by differing conditions of preservation of artefacts, and partly by the particular interests of each author or of Japanese archaeologists in general. For the Palaeolithic, stone tools predominate in the discussion; in the Jomon culture, pottery, subsistence and settlement types; in the Yayoi, early agriculture and bronze bells. The chapter on the Kofun period is inevitably devoted to a study of the development and structure of its huge keyhole-shaped mounds, as well as to evidence for social differences in the society. Finally, the descriptions of the early historical periods are concerned with texts, and sites related to the regional administrative bureaucracy.

Within these chapters, however, lie accounts of the discoveries that have placed Japan on the world archaeological map. First, the recognition (since 1949) of

the existence of its Palaeolithic cultures, and especially their remarkable features such as edge-ground axes (a technological achievement also found among the Pleistocene occupants of northern Australia), the oldest definite pottery vessels in the world (from 10,000 or even 12,000 years ago) and, since the book was written, evidence for very early sophistication in woodworking (see *Nature* 329, 110; 1987). In later periods came the establishment of rice cultivation, which has been investigated through excavation of paddy fields and analysis of plant phytoliths. And, in the early historical periods, the tens of thousands of slim wooden tablets bearing writing in ink, which pro-

vide a detailed picture of administration, economics and politics.

The great misfortune of Japanese archaeology, however, is that such a huge proportion of effort and resources over the past few decades has had to be devoted to rapid salvage excavations in the face of ever-encroaching development and construction projects. The result — the comparative lack of regional synthesis or planned research programmes — is all too evident in the book. It will be sad if studies of that kind can only be undertaken after most of Japan's sites have disappeared for ever. □

Paul G. Bahn, 428 Anlaby Road, Hull HU3 6QP, UK, is a freelance writer on archaeology.

Fun in the night

Bruce Margon

First Light: The Search for the Edge of the Universe. By Richard Preston. *Atlantic Monthly Press, New York*: 1987. Pp.263. \$18.95.

I HAD expected to dislike this book. The dust jacket (presumably not entirely the author's fault) is overripe in the extreme. Astronomers "rip the veil from space-time" but are not allowed to touch the controls of the huge 200-inch telescope "because they might destroy it". As an observational astronomer, I confess as well to some trepidation at an obvious although unstated purpose of this story — that is, finally to reveal how much fun a few grown men and women can have playing with computers and many tons of precision glass and steel in the middle of the night.

Despite my no doubt pompous intent to condemn this work, I failed. It's fun! Preston strings together a dozen essentially independent vignettes of astronomy in a somewhat odd, but certainly not unsatisfying, way. The book contains short biographies of a varied group of living and dead twentieth-century astronomers, sharing in common perhaps only their intense levels of energy, ingenuity and devotion to their work. Although the characters (such as Fritz Zwicky, Maarten Schmidt and Jim Gunn) are well known to other astrophysicists, they have not been previously featured in the general literature.

Interleaved between the personalities are tales of two current, important but again disparate astronomical problems: the quest for very high redshift quasi-stellar objects at the far edge of the Universe, and the search for a rare class of nearby asteroids at our cosmic doorstep. And for good measure, we learn something of the construction history of two major astronomical facilities, the Hale

200-inch telescope, and the Hubble Space Telescope.

Preston's style, although not quite as flowery as that on the dust jacket, will never be accused of being dry. The superlatives, expletives and adjectives fly past. But the author's enthusiasm is infectious, and the style soon becomes unobtrusive. The book flows back and forth between its tales smoothly, and with some of the urgency of a good mystery. The presentation is appropriate for an educated layman curious about twentieth-century science, but without a background in astronomy. Even professional astronomers and physicists will enjoy the colourful accounts of their famous colleagues' youth.

A more balanced and thereby more informative book might have at least briefly noted the parallel scientific activities of 'the competition'. The reader here may gain the impression that distant quasars are found only at Palomar Mountain, when in fact other groups around the world, particularly in Britain, were simultaneously using different but equally elegant search techniques, in some cases with spectacular successes that preceded those of the protagonists of *First Light*.

There are not enough permanent jobs in astrophysics to accommodate the number of bright new PhDs. There are certainly too few large telescopes as well. *First Light* is fun enough to read that it will do its part to make things worse. Books that, in a reasonably accurate and well-informed manner, show that scientists can make poor guesses, swear, trip on things and generally exhibit the full range of human failings, while still making genuine and lasting intellectual progress on the most difficult of problems, are important. This is such a book, and an interesting, entertaining and sometimes outright funny one at that. □

Bruce Margon, currently at the Anglo-Australian Observatory, PO Box 296, Epping, NSW 2121, Australia, is a Professor of Astronomy at the University of Washington.

Nuclear power

J.R.S. Fincham

Beyond the Gene: Cytoplasmic Inheritance and the Struggle for Authority in Genetics.

By Jan Sapp. Oxford University Press: 1987. Pp.266. \$35, £32.50.

GENETICISTS have always tended to believe that organisms are determined by genes located on chromosomes in the cell nucleus. Until the advent of gene cloning, this belief depended entirely on the Mendelian behaviour of nearly all variation that could be inherited through sexual crosses. Sceptics were able to argue that Mendelian variation might be only scratching the surface of an essentially invariant species-specific substructure. This was the view of some evolutionists, who thought that differences between species and higher-order taxa were based on something more fundamental than the chromosomes. Embryologists also tended to doubt the overriding importance of the chromosomes, whose apparent constancy made them seem quite unhelpful in accounting for development. Thus, during the great expansion of chromosomal genetics during the 1920s and 1930s, there was always some scepticism about genetics on the part of other biologists. At the same time, some geneticists began to

undermine the chromosomal monopoly in another way, by finding evidence for genetic determinants in the cytoplasm.

Dr Jan Sapp's history of these doubts and disputes makes interesting and provocative reading. The interest comes largely from extensive and well-selected quotations from the publications and (in some cases) the personal correspondence of some of the main protagonists. The provocation, at least for this reviewer, comes from the author's determination to explain the development of genetics in terms of struggles for authority and power between a Mendelian orthodoxy and cytoplasmic heresy.

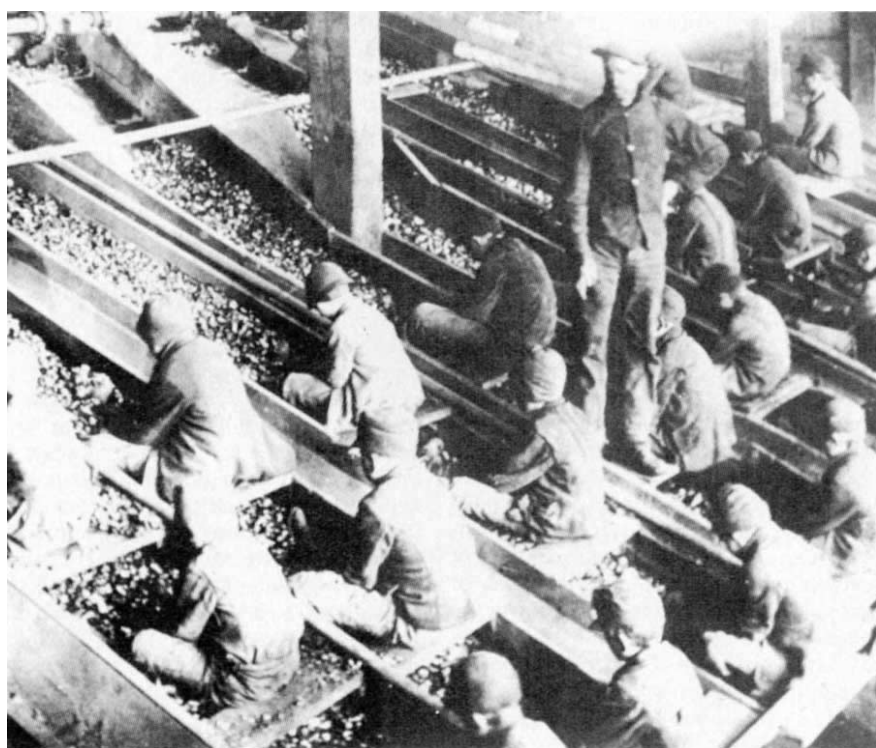
The three central chapters are mainly concerned, respectively, with the strong inter-war school of German extranuclear geneticists, the work of Tracy Sonneborn at the University of Indiana, and the post-war development of yeast (*Saccharomyces cerevisiae*) mitochondrial genetics in Paris under the leadership of Boris Ephrussi. In Germany, Carl Correns and his successors, notably von Wettstein, Renner and Michaelis, found many examples of maternal inheritance in green plants, mostly involving chloroplast function. This led them to postulate a genetic role for the plastids and to the concept of a 'plastome' acting in parallel with the genome. This was not an unorthodox view in Germany, in striking contrast to the United States where extranuclear

heredity received scant consideration. Dr Sapp attributes this to the different styles of science administration in the two countries; the American 'free market' pulled research into the most readily exploitable areas, whereas in the more professorially directed German system founder effects were of more importance. According to Dr Sapp, the American verdict on the German work was "not proven". But his quotations from American textbooks of the period show a general, if tentative, acceptance of a role for plastids in heredity, an acceptance encouraged, as he points out, by the fact that they were visible structures with apparent continuity through cell division.

The work started by Ephrussi led to the extension of genetic status from plastids to mitochondria. Although he had close links with the American *Drosophila* geneticists, especially A.H. Sturtevant and G.W. Beadle, Ephrussi was always sceptical about monopoly control by chromosomal genes — a legacy, the author thinks, of his background in embryology. His study of 'petite' (respiration-defective) mutants in yeast, later extended triumphantly to the DNA level by Piotr Slonimski in Gif, was largely responsible for extranuclear inheritance being welcomed into the mainstream of genetics. The early recognition of respiration deficiency as based in the mitochondria, followed by the firm identification of mitochondrial DNA, made it possible for Slonimski, and a number of other workers in strong laboratories in the United States and Australia, to explain many kinds of respiration defect as due to deletions or other mutations in specific mitochondrial genes. At nearly the same time, DNA was also identified in chloroplasts and a similar detailed analysis of chloroplast genes was soon under way. So far as this part of his subject is concerned, Dr Sapp's book might have been better entitled *The Gene Rides Again — Outside the Nucleus*.

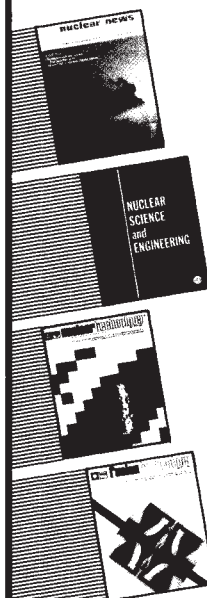
Sonneborn's experimental organism, the ciliate *Paramecium aurelia*, proved a rich source of examples of non-Mendelian inheritance. One of these, the killer phenotype, turned out to be due to endosymbiotic bacteria which had, of course, their own DNA. Other examples were not easily accommodated within the DNA paradigm. One of these concerned the specific surface patterning of the cell. Sonneborn showed that an alteration of the pattern by microsurgery could be transmitted through an indefinite number of subsequent cell divisions, an example of new structure guided by pre-existing structure, a mode of inheritance perhaps restricted to cells that never have to undergo cycles of dedifferentiation and redifferentiation.

Equally novel were the results of experiments carried out in collaboration



Back-breaking — children known as 'breaker boys' earned a dollar a week clearing the flow of pieces of coal through the chutes in an anthracite mine of the 1890s. The picture is taken from *American Environmental History*, 2nd edn, by Joseph M. Petulla, an illustrated social and economic history of the United States in the light of the nation's changing environment and of conservation issues. Publisher is Merrill, Columbus, Ohio, price is \$24.95.

The ANS Journals



Your number one source for current information about

NUCLEAR ENERGY

NUCLEAR NEWS

NUCLEAR SCIENCE & ENGINEERING

NUCLEAR TECHNOLOGY

TRANSACTIONS

REMOTE SYSTEMS TECHNOLOGY PROCEEDINGS

FUSION TECHNOLOGY

For free catalog contact:

AMERICAN NUCLEAR SOCIETY
555 N. Kensington Avenue
La Grange Park, Illinois 60525 USA
312-352-6611 Telex: 4972673

Reader Service No.22

THE MANAGEMENT OF AIDS PATIENTS

Edited by DAVID MILLER,
JONATHAN WEBER and JOHN GREEN

The Management of AIDS Patients is the first comprehensive guide to the practical clinical management of patients with AIDS or HTLV III infections. The book avoids the sensational aspects of the disease, offering solid advice and information for all people involved in patient care.

The editors and many of the contributors come from St Mary's Hospital, London one of the foremost centres in Britain for the treatment of AIDS patients. The knowledge and experience of these experts has been combined to provide a much-needed book for doctors, nurses, dentists and indeed for health-care professionals everywhere.

1986

200pp	Hardback	0 333 40465 3	£30.00
	Paperback	0 333 40466 1	£12.95

Please order this title from your bookseller – or in case of difficulty from Felicity Davie, Macmillan Press, Houndmills, Basingstoke, Hants RG21 2XS.

M MACMILLAN PRESS

with G.H. Beale on the inheritance of alternative antigenic specificities. The class of antigen to be expressed from among a number of alternatives was determined by the cytoplasm, with the possibility of temperature-induced switching from one class to another. But the fine specificity within each class was determined by alleles of a nuclear gene. Sonneborn postulated the existence of 'plasmagenes', replicating in the cytoplasm but drawing some information from the nucleus. An alternative interpretation was advanced by Max Delbrück, who suggested that cytoplasmic propagation of alternative patterns of antigen synthesis might be explained in terms of steady states, with gene products acting so as to stabilize the conditions for their own synthesis. This interpretation, interestingly enough, was considered as quite reasonable by Ephrussi in his book *Nucleocytoplasmic Relationships in Micro-organisms*, published in 1952. It still seems reasonable today, but the phenomenon has still not been adequately investigated at the molecular level. Sonneborn's discoveries remain highly interesting, but more as models for cellular differentiation than as challenges to gene determination.

Chapter 6, headed "The Cold War in Genetics", recounts, among other things, the leading part that Sonneborn took in trying to mobilize opinion against Lysenkoism in the Soviet Union; he happened to be President of the Genetics Society of America at the time. In spite of the occasionally expressed misconception that cytoplasmic heredity was Lysenkoist, Lysenkoism had nothing to do with the location within the cell of the genetic determinants. Lysenko denied the existence of genetic determinants as such and, moreover, he succeeded in getting this view of biology made compulsory in the Soviet Union. The clear matter of scientific principle was somewhat muddled by the concurrent McCarthyite hounding of American leftists in general. This chapter is rather marginal to the main theme of the book but it is full of interest, recalling a period now largely forgotten.

Today, most geneticists would say that the historical controversies and confusions reviewed in this book have in principle been resolved. Even the problem of development seems soluble in terms of genes and gene products interacting in the dimensions of space and time. The pace here has been set by the *Drosophila* developmentalists, who derived their initial inspiration largely from E.B. Lewis, a product of the Caltech school of genetics, represented in this book as the citadel of stifling orthodoxy. Dr Sapp, however, is not really interested in reconciliation of viewpoints. Chapter 7, "Problems with Master Molecules", is an attempt to show that the struggle

continues, though the principles at issue are not made altogether clear.

Throughout the book there is a steady tendency to under-rate the open-mindedness of geneticists. To take one example, in 1949 G.W. Beadle wrote a semi-popular article, "Genes and Biological Enigmas", which included a careful review of Sonneborn's main results. He concluded that there could be "no doubt whatsoever" about the existence in *Paramecium* and other organisms of cytoplasmic units "capable of a limited degree of autonomy"; but he went on to doubt whether they were responsible for "primary control of protein specificity". There are some interesting points to be made here; the lack of knowledge of protein structure and synthesis at the time Beadle wrote, and the tricky nature of the term 'autonomy'. But Dr Sapp seems to regard Beadle's article merely as a polemic; in his summarizing chapter he refers back to it as a "public attack" on Sonneborn's work.

Other examples cited in support of the book's thesis verge on triviality. Thus the journal *Genetics* once rejected a paper, nothing to do with cytoplasmic inheritance, apparently on the grounds that it was not based on adequately defined crosses. This becomes a "mechanism of control to maintain orthodoxy". Some correspondence between Ephrussi and Norman Horowitz, which most would regard as an exchange of banter between friends, is elevated to the status of "social negotiations" about adjustments in authority. In his final chapter Dr Sapp resorts increasingly to arcane sociological jargon to describe interactions between people. I interpret this as part of the struggle for authority in history-of-science writing.

To be fair, Dr Sapp seems not totally opposed to the view of scientists as seekers after true knowledge. Although in his final chapter, he gives the "predominant role" to "power relationships within and among scientific disciplines", he still, one is relieved to read, assumes the existence of a material reality that imposes some constraints on the outcome of scientific controversies. Nor does he discount the achievements of molecular biology. Only once or twice does he stray into anti-reductionism, that quagmire of confusion. The view of scientific progress expressed in this book is, in my view, a distorted one, giving insufficient credit to the open-mindedness and mutual respect that usually (but not always) attends the scientific endeavour. But there is a lot of fascinating material here on the careers of important scientists and the ideas with which they struggled. □

J.R.S. Fincham is Balfour Professor in the Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK.

A better way to control pollution

Dick Derwent

Pollutants emitted into the air can damage the environment far from their original sources. Control strategies can best be formulated through use of optimization techniques.

OPTIMIZATION by simulated annealing is a novel technique which has recently been applied to some seemingly intractable problems¹ in such diverse disciplines as microelectronic circuit design, solid-state physics, design of neural networks and operational research. The long-range transport of air pollutants is a further example of an area in which the technique can fruitfully be used to provide quantitative information. Given enough background data, one can compare environmental protection strategies which have specific objectives of reducing the effects of pollution, taking into account both practical and economic constraints on pollution-control technology. The example used here is the long-range transport of nitrogen oxides (NO_x).

Within Europe, there is strong international pressure to control the emissions of nitrogen oxides from power stations and motor vehicle exhausts, and so reduce acidification of the environment. These pollution sources give rise to nitric oxide (NO), which is converted in the atmosphere to nitrogen dioxide (NO_2), nitric acid (HNO_3) and nitrate aerosol (pNO_3), which may be transported and then deposited up to 1,000 km from their original sources.

Measurements² of the nitrate to sulphate ratio in the atmospheric aerosol in southern England have shown a steady increase since 1954. The nitrate content of precipitation averaged over the entire European Air Chemistry Network³ has steadily increased over the period 1955 to 1979. The nitrate levels in ice cores from South Greenland have continued to increase steeply from 1975 to 1984, whilst sulphate has remained relatively constant since 1968⁴. The 'Thousand Lake Survey' in Norway⁵ has recently revealed a doubling in the nitrate concentration of 305 lakes over the period 1974–1975 to 1986, despite little change in pH and sulphate.

Formulation of a coherent NO_x control strategy for Europe must be based on an understanding of the life cycle of NO_x in the environment. But the analysis described here shows that reduction of NO_x emissions by a fixed percentage in all European countries is not the best approach, and that other strategies offer the possibility of more dramatic reductions in deposition for similar levels of investment in pollution control. The most cost-effective way forward will involve

close cooperation between two distinct groups of European countries.

The formulation of pollution control strategies depends upon theoretical modelling studies⁶ to unravel the processes which link NO emissions to acidic nitrogen deposition in Europe. It further requires the selection of sensitive deposition (receptor) sites, the assessment of critical deposition loads at these sites and the application of an optimization methodology which has the power and flexibility to cope with the detailed aspects of regional-scale pollution problems. At each step in this study, examples are used to develop the complete framework and provide some illustrative results. When environmental research has furnished the necessary criteria, these assumptions will be improved both spatially and quantitatively.

Table 1 lists the 11 receptor sites that have been chosen to illustrate the methodology. In the absence of means to identify places that are particularly sensitive to man-made acidic nitrogen species, the sites have been selected solely to illustrate the analysis. They are, however, subject to acidic nitrogen deposition and possibly to environmental damage^{7,8} caused by acidification of surface waters and aquatic ecosystems, acidification of sensitive soils, eutrophication of marine environments, and nitrogen saturation effects in forests, heathlands and bogs. Dry and wet deposition are assumed to be equally damaging. When target or critical loads become available then these could readily be taken into account.

Our understanding of the atmospheric behaviour of acidic nitrogen species has been much enhanced by theoretical studies. A simple trajectory model⁶ is

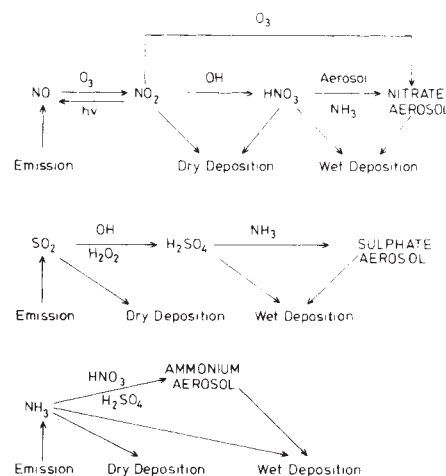


Fig. 1 Coupled life cycles of the sulphur- and nitrogen-containing trace gases.

used here to relate the nitric oxide emissions on the United Nations Economic Commission for Europe (UNECE) European Monitoring and Evaluation Programme (EMEP) 150 km × 150 km grid to acidic nitrogen deposition at each of the receptor sites. In this model, an air parcel extending from the ground to the top of the boundary layer is advected by the wind over the emissions grid to reach the receptor site. A single trajectory layer has been adopted and instantaneous mixing within the air parcel is assumed. The pollutants within the air parcel are subject to chemical transformation involving the coupled chemistry of NO_x , SO_2 , NH_3 and ozone (see Fig. 1, the details of which are presented elsewhere⁹), before deposition.

It would have consumed too much computer time to use the model to evaluate the deposition at each receptor site for each NO_x control strategy, and so a sim-

Table 1 Receptor sites possibly sensitive to the deposition of acidic nitrogen species

Site	Latitude*	Longitude*	EMEP co-ordinates	Nearest EMEP site†
Central Wales	52°20'N	2°40'W	16,13	Ludlow
South-west Scotland	54°40'N	5°00'W	15,15	Eskdalemuir
The Netherlands	52°40'N	5°40'E	20,16	Witteveen
West Norway	60°40'N	4°40'E	16,21	Fitjar
South-west FRG	48°20'N	7°20'E	23,14	Schauinsland
South Norway	58°00'N	6°40'E	18,20	Birkenes
North Sea	54°20'N	7°40'E	20,18	Westerland
South-east FRG	49°00'N	12°00'E	24,16	Brotjackriegel
South-west Sweden	57°40'N	12°20'E	20,21	Rorvik
Baltic Sea	55°20'N	15°40'E	22,20	Duodde
South Finland	59°40'N	23°20'E	21,25	Uto

*To the nearest ± 20'.

†See ref. 13.

pler approach was taken which nevertheless retains the main aspects of the NO_x life cycle. A nitric oxide source was set up in a single $150 \text{ km} \times 150 \text{ km}$ grid square in the centre of Europe and the total nitrate deposition was assessed at each EMEP grid line intersection using the trajectory model. The resulting spatial deposition pattern defines the source-receptor relationship (Fig. 2). A simple analytical expression adequately represents the relationship found between NO_x emission and total deposition at a given distance away. Estimates of total deposition at each receptor site were obtained by summing the contributions from every grid square, with each contribution proportional to the grid square emission and scaled by the source-receptor relationship in Fig. 2.

The pattern of NO_x emissions across Europe on the $150 \text{ km} \times 150 \text{ km}$ baseline grid for 1985 (ref. 10) provides the basis of this study. To represent an NO_x control strategy, it is imagined that a pollution-abatement device, with a fixed capacity to reduce NO_x emissions by $10 \text{ ktonne yr}^{-1}$, is installed in a particular grid square. Any number of these fixed-capacity devices can be installed up to the total reduction in NO_x emissions required, and they may be distributed in any spatial pattern over the grid. A control strategy is designed by prescribing the spatial pattern and density of the devices. By subtracting the pollution-control capacity from the 'baseline' 1985 emissions for each grid square, a new 'controlled' emission grid is generated; hence, new values for the total deposition at each of the receptor sites can be calculated either with the trajectory model or using the simple source-receptor relationship described above.

The simplest NO_x control strategy is that of 'blanket percentage reductions in emissions', which calls for a fixed percentage reduction in emissions in each grid square. To illustrate this, the emissions in each grid square were reduced by 30 per cent. As a result, the total deposition at each of the 11 sites decreased by between 25.8 per cent and 29.6 per cent, reflecting the inherent linearity in the NO_x system and the negligibly small background nitrate deposition⁶. This strategy therefore gives roughly the same percentage reduction in deposition at all receptor sites, but it involves installing pollution-control equipment in grid squares which are far upwind of the sensitive receptor sites and which thus contribute little or nothing to reducing acidification of the environment there.

More cost-effective NO_x control strategies were investigated using optimization by simulated annealing¹¹, which has not been previously applied to environmental problems. The technique involves iterative improvement, in which the system starts in a known configuration and a standard but random rearrangement is applied. If

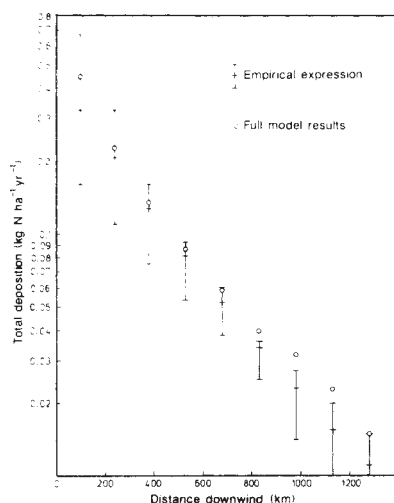


Fig. 2 Comparison of the long-term average total acidic nitrogen deposition at a point downwind of an NO_x source, emitting 100 ktonne NO_2 per year, in an isolated EMEP grid square.

the new configuration represents a net improvement on the original then it is kept and the random rearrangement applied again. This process is repeated until there are no further improvements, and the method usually gets stuck in a local deposition minimum rather than the required global minimum. Optimization by simulated annealing employs this method with a powerful extension based on the Metropolis algorithm¹², which comes from statistical mechanics and provides a generalization of iterative improvement in which controlled uphill steps can be incorporated in the search for a better solution.

The present implementation has the locations of the constant-capacity abate-

ment devices as the configuration of the system, with each abatement device subject to a random walk over the EMEP grid. An objective function can be defined in terms of the deposition at one of the receptor sites from Table 1 or some combination of up to all 11 of them. For a given total reduction in NO_x emission, simulated annealing can determine the optimal siting of the abatement devices to minimize deposition as defined through the objective function. The detailed implementation of the optimization techniques is described elsewhere⁹.

Taking each of the 11 receptor sites in turn, the optimal siting was found for abatement devices able to reduce NO_2 emissions by an annual total of one million tonnes. Table 2 shows that the reductions in deposition were far greater than the 6.7 per cent expected from the strategy of blanket percentage reductions in emissions involving equipment of the same total abatement capacity and hence the same costs. Importantly, optimal strategies could be found for all the receptor sites taken separately, and it was possible to study the siting patterns to find out why the optimal strategies were so efficient. In all these calculations, it is assumed that each reduction of 10 ktonne of NO_2 per year has the same cost irrespective of location in the grid. This implies that the costs of emission-control technology are independent of country and that the costs of pollution abatement are independent of extent of control. The 'last' 10 ktonne of reduction in emissions may well cost a lot more than the 'first' 10 ktonne , but there is not yet enough international information to construct a cost algorithm which addresses this point.

Figure 3 shows how simulated annealing was able to identify optimal NO_x

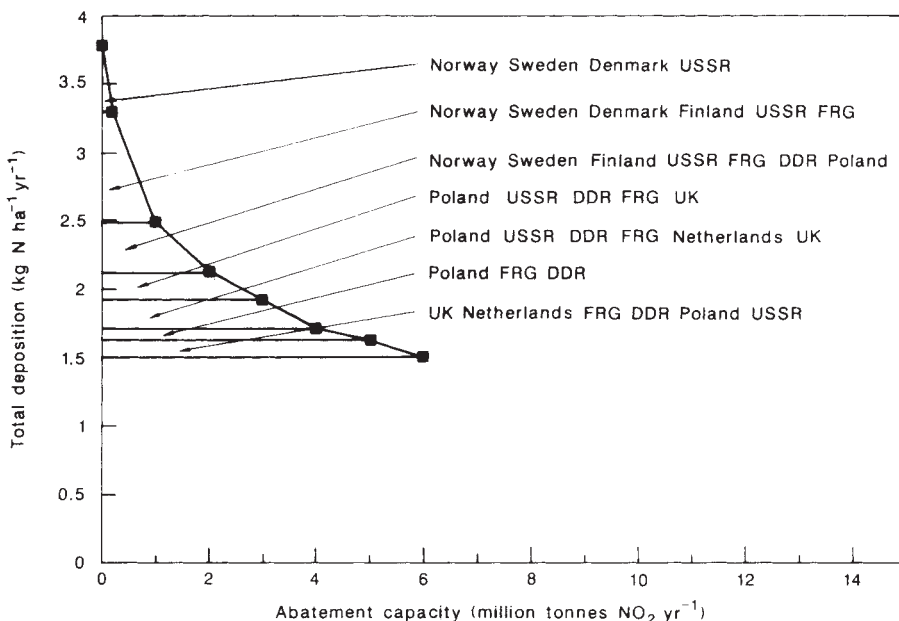


Fig. 3 Variation in total acidic nitrogen deposition in south-west Sweden for increasing capacities of NO_x abatement, sited using simulated annealing.

control strategies at one receptor site in south-west Sweden for increasing total capacities of NO_x abatement. In each optimal strategy, pollution-abatement devices are installed progressively from the receptor site outwards. The shape of the curve of total deposition against installed abatement capacity exhibits the classic behaviour of the law of diminishing returns. Each progressive tranche of abatement capacity has less and less of an effect on reduction of deposition levels. The most cost-effective control of emissions is achieved by installing the first tranche of abatement capacity close to the receptor site, with subsequent tranches installed further and further away. This behaviour is a direct consequence of the emission-deposition relationship derived from the life-cycle behaviour of the acidic nitrogen species as formulated in the trajectory model. Diminishing returns and the advantages of installing pollution-control equipment closest to the receptor site are seen at all the receptor sites listed in Table 1.

When all the receptor sites are taken together, optimization techniques seek the spatial pattern of NO_x abatement which minimizes deposition at all the sites. Here it is important to learn how the trade-offs work on the European scale. Each fixed-capacity abatement device, when installed in a particular grid square, reduces the deposition differently at each of the receptor sites and hence decreases the objective function. As it is randomly moved across the grid in the rearrangement process, the inherent emission-deposition relationship determines how the deposition increases at one site yet decreases at another and so on. When the objective function, and hence deposition at all the sites, is at a minimum then all the trade-offs have been resolved and the pollution-abatement device is at its optimal position on the grid.

Table 3 shows the reductions in deposition achieved at each receptor site in an optimal strategy for the installation of devices able to reduce NO_2 emissions by one million tonnes a year. Optimization by simulated annealing was able to reduce simultaneously the deposition at 9 of the 11 receptor sites by much more than the 6.7 per cent expected from blanket percentage reductions in emissions. At the remaining two sites, there were only small reductions in deposition. Whatever objective function I have tried, I have been unable to find an optimal strategy which encompasses all 11 sites.

From a policy point of view, there are two distinct geographical areas which show different responses to NO_x controls. The first takes in the receptor sites on the fringes of the European continent, the second those in central Europe. The results of simulated annealing dictate separate strategies for these two areas, both involv-

Table 2 Total deposition ($\text{kg N ha}^{-1}\text{yr}^{-1}$) at each receptor site in the base case and in an optimized strategy, each site considered individually*

Site	Base case	Optimized strategy (1.0 million tonnes $\text{NO}_2 \text{ yr}^{-1}$ abatement)	Percentage decrease on base case
Central Wales	4.614	2.358	48.9
South-west Scotland	2.536	1.345	46.9
The Netherlands	8.963	6.698	25.3
West Norway	1.324	0.991	25.2
South-west FRG	15.460	11.243	27.3
South Norway	2.755	2.005	27.2
North Sea	5.988	4.324	27.8
South-east FRG	14.368	10.017	30.3
South-west Sweden	3.804	2.495	34.4
Baltic Sea	5.259	3.540	32.7
South Finland	2.176	1.026	52.8

Table 3 Total deposition ($\text{kg N ha}^{-1}\text{yr}^{-1}$) at each receptor site in the base case and in an optimized strategy, receptor sites considered collectively*

Site	Base case	Optimized strategy (1.0 million tonnes $\text{NO}_2 \text{ yr}^{-1}$ abatement)	Percentage decrease on base case
Central Wales	4.614	4.331	6.1
South-west Scotland	2.536	2.121	16.4
The Netherlands	8.963	8.054	10.1
West Norway	1.324	1.099	17.0
South-west FRG	15.460	14.706	4.9
South Norway	2.755	2.253	18.2
North Sea	5.988	4.812	19.6
South-east FRG	14.368	12.353	14.0
South-west Sweden	3.804	2.955	22.3
Baltic Sea	5.259	4.133	21.4
South Finland	2.176	1.811	16.8

*In both Table 2 and Table 3 deposition estimates are quoted to three figures not because of inherent precision but to highlight small differences. Total deposition includes contributions from dry deposition of NO_2 , HNO_3 and pNO_3 , and wet deposition of HNO_3 and pNO_3 .

ing local control ahead of controls in more distant sources depending on how much of a reduction in deposition is required. The difference lies in the grouping of the countries between which co-operation is required to secure any required reductions in deposition. The two groups are Central Europe: Federal Republic of Germany (FRG); German Democratic Republic (DDR); Switzerland; and France. Fringe of Europe: Scandinavia; United Kingdom; Belgium; Netherlands; Poland; FRG; DDR; and USSR.

The optimized patterns of emission reduction found by simulated annealing are dramatically different, as Fig. 4 (overleaf) shows, depending on whether maximum reductions in deposition are sought in central Europe or the fringe sites. For the central Europe sites, abatement capacity should be concentrated in the FRG and its neighbouring countries, whereas for the fringe sites NO_x abatement technology is best installed over a wide area and in many countries.

An increasingly important forum for hammering out the policy approach to regional-scale air pollution problems is the international Convention on Long Range Transboundary Air Pollution (LRTAP), under the auspices of the UNECE. The convention has also become

a driving force behind the research programmes investigating the underlying science of long-range transboundary pollution. Sulphur dioxide was the first pollutant to be discussed within this framework. The protocol that resulted binds all signatories to reduce, by 1993, annual SO_2 emissions by 30 per cent from their 1980 level, a commitment based on pragmatism rather than science. The research programmes within the convention are now turning to other emerging regional-scale pollution problems, including nitrogen deposition and photochemical ozone. These issues could lead ultimately to the formulation of new protocols under the LRTAP Convention.

Simulated annealing indicates that it may well be possible to control pollution much more effectively and more cheaply than could be achieved by a protocol expressed in terms of a blanket percentage reductions in emissions. However, the real situation is much more complex than I have been able to represent in this example. The structural differences in emissions between countries may prove to be particularly significant. But, with optimization by simulated annealing, a start has been made on NO_x , and further work will consider sulphur deposition and photochemical ozone. ►

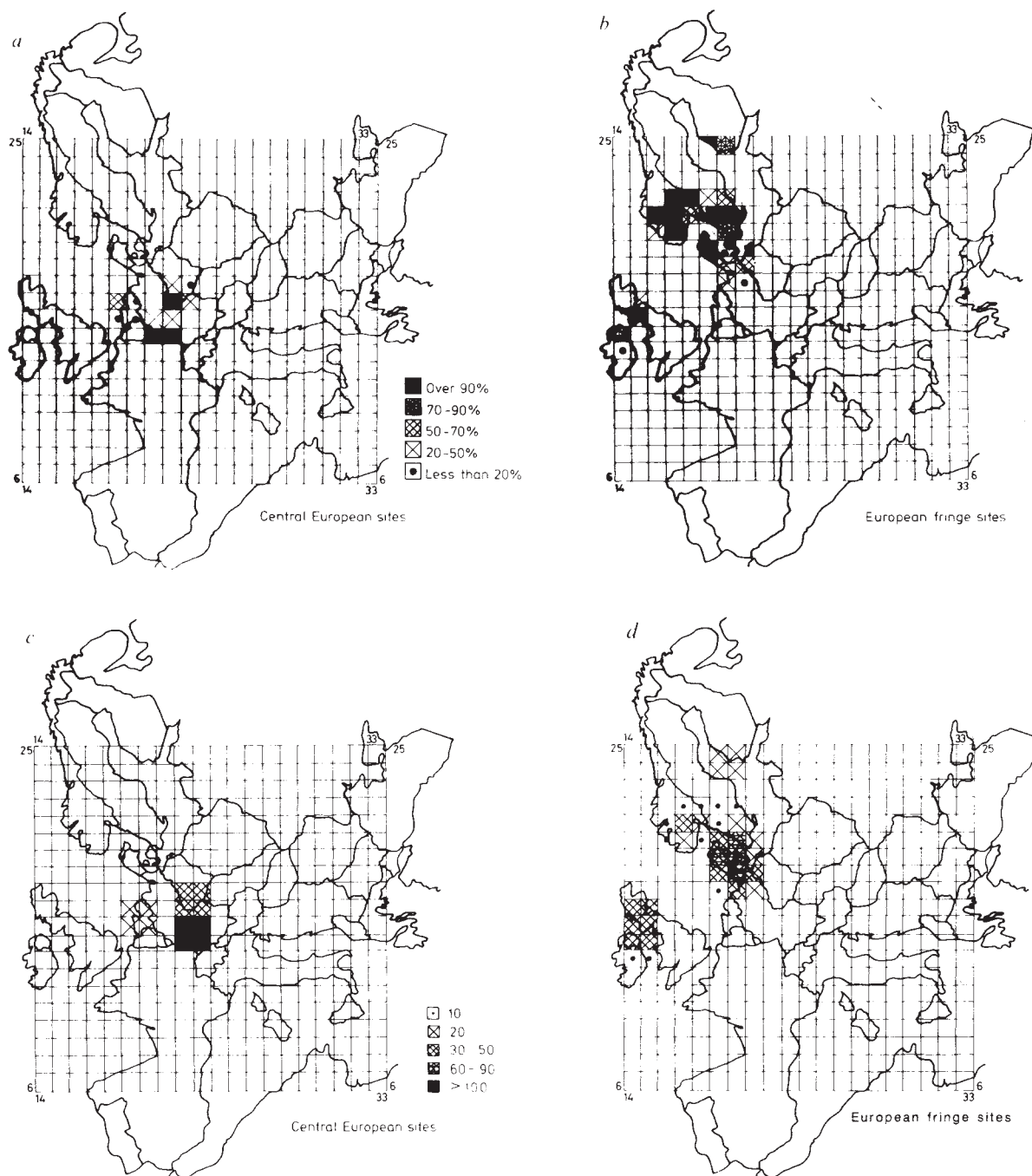


Fig. 4 *a, b* Percentage reduction in NO_x emissions required by simulated annealing to give the maximum reduction in deposition at the central European and the fringe sites. *c, d* Spatial pattern of NO_x emission-abatement required to give the maximum reduction in deposition at the central European and the fringe sites; figures are thousands of tonnes NO₂ yr⁻¹ of abatement capacity. Total emission-abatement capacity in all cases is one million tonnes NO₂ yr⁻¹.

This article is based on research funded by the UK Department of the Environment air pollution research programme. The views expressed are those of the author and are not necessarily those of the Department of the Environment.

1. Bounds, D.G. *Nature* **329**, 215-249 (1987).
2. Atkins, D.H.F. & Law, D.V. *Retrospective Trend Analysis of Atmospheric Sulphate and Nitrate at a Site in the United Kingdom Between 1954 and 1982* AERE Rep. R-11144 (Harwell Laboratory, Oxfordshire, 1984).
3. Soderlund, R., Granat, L. & Rodhe, H. *Nitrate in Precipitation* — A Presentation of Data from the European Air Chemistry Network (EACN) (Dept Meteorology, University of Stockholm, Stockholm, 1985).

4. Mayewski, P.A. *et al. Science* **232**, 975-977 (1986).
5. Bergmann-Paulsen, B. & Kvaeven, B. (eds) *1000-Lakes Investigation 1986* Statlig program for forurensningsovervakning, SFT Rep. 282/87 (Statens Forurensningstilsyn, Oslo, 1987).
6. Derwent, R.G. & Nodop, K. *Nature* **324**, 356-358 (1986).
7. Aniansson, B. (ed.) *Europe's Air Environment* (Norstedts Tryckeri, Stockholm, 1986).
8. Nilsson, J. *Critical Loads for Sulphur and Nitrogen* Nordisk ministerrad, miljo rapport 1986 No. 11 (Gotab, Stockholm, 1986).
9. Derwent, R.G. *Combating the Long Range Transport and Deposition of Acidic Nitrogen Species in Europe* AERE Rep. R-12799 (HMSO, London, 1987).

10. Eliassen, A. & Saltbones, J. *Emissions of NO_x in Europe* (Norske Meteorologiske Institutt, Blindern, Oslo, 1987).
11. Kirkpatrick, S., Gelatt, C.D. & Vecchi, M.P. *Science* **220**, 671-680 (1983).
12. Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. & Teller, E. *J.chem.Phys.* **21**, 1087-1092 (1953).
13. Schaig, J. *et al. Summary Report from the Chemical Coordinating Centre for the Third Phase of EMEP* EMEP-CCC Rep. 3/87 (NILU, Lillestrom, 1987).

Dick Derwent is in the Modelling and Assessment Group, Environmental and Medical Sciences Division, UKAEA Harwell Laboratory, Oxfordshire OX11 0RA, UK.

Manganese–chromium isotope systematics and the development of the early Solar System

Jean-Louis Birck & Claude J. Allègre

Laboratoire de Géochimie et Cosmochimie, Institut de Physique du Globe et Département des Sciences de la Terre, Université de Paris VI et VII, 4, Place Jussieu, 75252 Paris Cedex 05, France

High-precision measurements of the chromium isotope compositions of samples from meteorites reveal anomalies in the $^{53}\text{Cr}/^{52}\text{Cr}$ ratio which are believed to arise from in situ decay of the extinct short-lived nuclide ^{53}Mn . The decay of ^{53}Mn to ^{53}Cr in the early Solar System provides an additional chronometer with which to constrain the formation times of the small planetary bodies from which the meteorites originated. A comparison of chromium and titanium isotope anomalies shows them to be imperfectly correlated, bearing witness to the complexity of early Solar System processes.

THE discovery of isotopic variations in meteoritic material^{1,2} has led to significant changes in the models previously established for the origin of the Solar System. In particular, the presence of short-lived nuclides in the early Solar System has led astrophysicists to reconsider some of the well established interpretations of the record of nucleosynthesis preserved in early Solar System materials^{3–7}.

The iron-group elements have received special attention, because they represent the most stable nuclei in the periodic table and are synthesized in the ultimate stages of stellar evolution^{8–11}. Titanium and calcium isotope anomalies are widespread in inclusions in the Allende meteorite^{12–16}. Recently, chromium isotope anomalies¹⁷ have also been documented, revealing the presence of ^{53}Mn in the early Solar System^{18,19}.

In this paper, we present high-precision data demonstrating *in situ* decay of ^{53}Mn in early Solar System objects. Using ^{53}Mn – ^{53}Cr systematics, new light is shed on the chronology of planetary objects. We discuss also the anomalies observed for ^{50}Ti and ^{54}Cr , which are the most neutron-rich isotopes of these two elements, and the constraints that may be drawn from these observations for the formation of the host meteorites.

^{53}Mn – ^{53}Cr internal systematics

^{53}Mn decays to ^{53}Cr with a half-life of 3.7 Myr. Although now extinct, its one-time presence in early Solar System materials is revealed in variations in the $^{53}\text{Cr}/^{52}\text{Cr}$ ratio. The chromium–manganese conservation law for a closed system including ^{53}Mn is as follows:

$$\begin{aligned} \left(\frac{^{53}\text{Cr}}{^{52}\text{Cr}}\right)_{\text{today}} &= \left(\frac{^{53}\text{Cr}}{^{52}\text{Cr}}\right)_{\text{initial}} + \left(\frac{^{53}\text{Mn}}{^{52}\text{Cr}}\right)_{\text{initial}} \\ &= \left(\frac{^{53}\text{Cr}}{^{52}\text{Cr}}\right)_{\text{initial}} + \left(\frac{^{55}\text{Mn}}{^{52}\text{Cr}}\right)_{\text{today}} \left(\frac{^{53}\text{Mn}}{^{55}\text{Mn}}\right)_{\text{initial}} \end{aligned} \quad (1)$$

For a system having homogeneous initial ($^{53}\text{Mn}/^{55}\text{Mn}$) and ($^{53}\text{Cr}/^{52}\text{Cr}$) ratios, the representative ratios of various samples of this system fall on a straight line in the ($^{55}\text{Mn}/^{52}\text{Cr}$)–($^{53}\text{Cr}/^{52}\text{Cr}$) diagram. The slope of this line gives the $^{53}\text{Mn}/^{55}\text{Mn}$ ratio at the time of formation of the meteoritic object, and the intercept the initial $^{53}\text{Cr}/^{52}\text{Cr}$ ratio at the time of closure of the system. In an attempt to use this method, first used for the ^{129}I – ^{129}Xe couple²⁰ and then applied to the ^{26}Al – ^{26}Mg (ref. 21) and ^{107}Pd – ^{107}Ag (ref. 5) systems, we measured the (Mn/Cr)–($^{53}\text{Cr}/^{52}\text{Cr}$) couple in a number of meteoritic samples.

Various types of meteorite were investigated: the Allende meteorite and its refractory inclusions; the C2 meteorite Murchison; the LL3 Chainpur meteorite, in which hydrogen and

carbon isotope anomalies had already been found²²; the diogenite Tatahouine; the pallasite Eagle Station, in which an oxygen isotope anomaly has been found²³; and the (E4) enstatite-bearing chondrite Indarch, which contains high-Mn/Cr sulphides. Basalts from both the Earth and the Moon were analysed for the sake of completeness and testing.

A precise experimental technique has been developed to resolve the small effects reported here (see Table 1 legend). The data are shown in Table 1. Our results show that in all meteorite samples except Tatahouine and the chondrules of Chainpur, there is clear evidence of variation of $^{53}\text{Cr}/^{52}\text{Cr}$ in coexisting phases, correlated with Mn/Cr. When this evidence is obtained by preferential leaching of a bulk sample, the inferred $^{53}\text{Mn}/^{55}\text{Mn}$ ratio may be uncertain owing to the possibility of differential leaching for Mn and Cr, but we believe that this cause of error is of minor importance here, as for ^{26}Al and ^{26}Mg in Allende inclusions²¹.

The results can be summarized as follows:

Allende. Previously reported data¹⁸ on whole-rock Allende inclusions and for minerals from the coarse-grained inclusion BR1 are also included in Table 1. The Mn–Cr mineral data of BR1 are in agreement with the whole-rock data. The mean values for both sets of samples are $^{53}\text{Mn}/^{55}\text{Mn} = (4.4 \pm 1.0) \times 10^{-5}$ and $\epsilon^{53}\text{Cr} = -1.7 \pm 0.3$. (The ϵ notation is defined in the Table 1 legend.) A similar experiment was carried out on samples of fine-grained inclusions G1 and G6; two high-precision data points were obtained for each inclusion. With so few data we consider it premature to draw any conclusion about the Mn–Cr system in fine-grained inclusions, especially as the ^{26}Al – ^{26}Mg system displays complex features in fine-grained inclusions²⁴.

Indarch. In the enstatite chondrite Indarch, manganese is present in alabandite (MnS) and chromium in daubreelite (Cr_2FeS_4). These sulphide phases were dissolved with solvents of increasing acidity, alabandite being more readily soluble than daubreelite. Except during the H_2O leach, special care was taken to keep the pH below 4, to avoid the precipitation of chromium hydroxide. The analytical results show a good correlation (Fig. 1), yielding a $^{53}\text{Mn}/^{55}\text{Mn}$ ratio of 1.1×10^{-6} and an intercept value of $\epsilon^{53}\text{Cr} \approx 1$.

Eagle Station. For the Eagle Station pallasite (Fig. 2), a pure olivine fraction was picked mechanically. The fine-grained chromites are usually embedded in centimetre-size olivine crystals. A pure chromite fraction was obtained by dissolving the olivine from a selection of chromite-rich olivine crystals in diluted $\text{HF} + \text{HCl}$. The difference in $^{53}\text{Cr}/^{52}\text{Cr}$ ratio between the two points is more than eight times the error bars. The $^{53}\text{Mn}/^{55}\text{Mn}$ ratio is found to have been $(2.3 \pm 0.3) \times 10^{-6}$, with normal initial $^{53}\text{Cr}/^{52}\text{Cr}$.

Table 1 Chromium isotope compositions and Cr and Mn contents of the samples

Sample	Cr (p.p.m.)	Mn (p.p.m.)	$\epsilon^{53}\text{Cr}^*$	$\epsilon^{54}\text{Cr}^*$
Allende inclusions				
G1†	1,434	261	-1.08 ± 0.33	7.7 ± 0.6
G1 K silicate	834	242	-0.89 ± 0.16	7.0 ± 0.3
‡			-0.83 ± 0.19	8.2 ± 0.4
G1 K spinel	283	27.9	-0.93 ± 0.17	7.6 ± 0.3
‡			-1.11 ± 0.1	8.4 ± 0.2
G2 A†			-0.46 ± 0.80	6.6 ± 1.1
G2 B†	378	76§	-1.23 ± 0.50	5.9 ± 1.3
‡			-0.8 ± 0.58	5.0 ± 1.4
G6 silicate	643	372	1.40 ± 0.20	3.5 ± 0.5
G6 D	431	238.2	0.80 ± 0.17	3.3 ± 0.3
‡			1.03 ± 0.17	3.4 ± 0.4
G6 D spinel	144	46.4	0.64 ± 0.13	3.4 ± 0.3
‡			0.85 ± 0.14	4.0 ± 0.3
G4†	400	92	-0.47 ± 0.45	6.1 ± 1.0
SB	492	94§	0.00 ± 0.36	6.4 ± 0.7
BR9	383	47.7	-1.06 ± 0.36	7.7 ± 0.5
BR9 G silicate	225	20.3	-1.35 ± 0.26	6.1 ± 0.5
BR9 G spinel	170	25.5	-0.99 ± 0.16	6.9 ± 0.3
‡			-0.74 ± 0.14	7.0 ± 0.3
BR1 B†	381	86	-0.86 ± 0.73	4.9 ± 1.1
$\rho < 2.9$ melilite†	81.5	35	0.01 ± 1.8	2.4 ± 4.3
$\rho > 3.3$ pyroxene†	189	82	0.29 ± 0.74	6.1 ± 1.5
‡			-0.14 ± 0.44	6.1 ± 1.0
‡			0.30 ± 0.59	5.8 ± 1.3
‡			-0.19 ± 0.18	5.0 ± 0.4
$\rho > 3.3$ spinel†	293	19	-1.14 ± 0.37	6.2 ± 0.6
‡			-1.26 ± 0.23	6.1 ± 0.5
‡			-1.65 ± 0.62	5.0 ± 1.1
Allende bulk†	3,500#	1,490#	0.14 ± 0.26	1.7 ± 0.4
‡			0.46 ± 0.44	2.3 ± 0.9
HF-insoluble residue			-0.21 ± 0.64	3.4 ± 1.1
‡			-0.57 ± 0.41	2.6 ± 0.8
Murchison $\rho < 2.9$			0.67 ± 0.29	3.2 ± 0.6
Tatahouine			0.16 ± 0.32	0.2 ± 0.6
Chainpur chondrule 1			0.25 ± 0.39	1.1 ± 0.6
‡			0.46 ± 0.71	1.9 ± 1.5
chondrule 2			0.18 ± 0.47	-0.2 ± 0.8
chondrule 3			0.10 ± 0.63	-0.9 ± 1.3
‡			0.29 ± 0.42	-0.0 ± 0.8
Eagle Station olivine	180	1,320	1.74 ± 0.46	2.4 ± 1.2
‡			1.79 ± 0.18	1.6 ± 0.4
‡			1.57 ± 0.15	1.2 ± 0.3
chromite	43.5	0.408	-0.05 ± 0.15	1.5 ± 0.3
			0.12 ± 0.19	1.5 ± 0.4
Indarch H ₂ O leach	0.09	13	< 10	< 10
Dilute acetic acid	14.87	240	2.88 ± 0.43	0.2 ± 0.9
‡			2.65 ± 0.43	0.9 ± 0.9
‡			2.88 ± 0.49	0.2 ± 1.0
50% acetic acid	12.04	44	1.56 ± 0.67	1.0 ± 1.2
‡			1.63 ± 0.46	0.7 ± 0.9
37% HNO ₂	1,926	1,018	0.98 ± 0.43	1.8 ± 0.7
‡			0.74 ± 0.83	1.2 ± 1.4
‡			0.67 ± 0.81	2.3 ± 1.4
Silicate residue	< 200		0.25 ± 0.33	0.8 ± 0.7
Tibet ophiolite			-0.19 ± 1.0	0.0 ± 1.8
Lunar 74261,1 soil			-0.27 ± 0.73	-0.5 ± 1.5

After dissolution of the sample, the solution is converted to the chloride form. This separation method makes use of the slow re-equilibration of the different hydrated species of Cr^{3+} in aqueous solution after a change in the concentration of the counter-ions. Two successive elutions in 1-N HCl are performed on a Biorad AG 50WX8 cation exchanger. We developed a mass spectrometric method using silica gel and boric acid as a high-yield emitter for solid-source mass spectrometry. Interferences from ^{50}Ti , ^{50}V and ^{54}Fe were always negligible. Chromium amounts of ≥ 300 ng yield precisions of 2×10^{-5} on the $^{53}\text{Cr}/^{52}\text{Cr}$ ratio and $(5-8) \times 10^{-5}$ on $^{54}\text{Cr}/^{52}\text{Cr}$. Although some data are obtained with standard deviation significantly lower than these values and display good self-consistency, these values are used as lower limits for the errors in any further calculations using the data (slope and intercept), as we think that these values represent the current limits of our instrument. Cr contents were determined by isotope dilution, with experimental precision better than 1%. Mn contents were measured by atomic absorption with and without a graphite furnace; the uncertainty is $< 5\%$ unless specified. The $^{55}\text{Mn}/^{52}\text{Cr}$ ratio plotted in the figures is derived from the elemental ratio by multiplying by 1.130.

* Values and errors (95% c.l.) are expressed as ϵ values relative to the mean value of the laboratory standard, where ϵ is defined as $10^4 \times (\text{ratio measured} - \text{standard ratio}) / \text{standard ratio}$. $\epsilon^{53}\text{Cr}$ refers to the $^{53}\text{Cr}/^{52}\text{Cr}$ ratio; $\epsilon^{54}\text{Cr}$ refers to $^{54}\text{Cr}/^{52}\text{Cr}$. The NBS value⁵⁰ of $^{52}\text{Cr}/^{50}\text{Cr} = 19.28323$ is used as a normalization value for instrumental fractionation correction. The grand mean values for the laboratory standard are: $^{53}\text{Cr}/^{52}\text{Cr} = 0.1134569 \pm 0.0000010$; $^{54}\text{Cr}/^{52}\text{Cr} = 0.02821008 \pm 0.0000005$. For Allende inclusions the term 'silicate' is applied to phases soluble in $\text{HF} + \text{HNO}_3$, and 'spinel' to phases insoluble in this medium. For the samples on which selective dissolution was carried out (G1, G6, BR1, BR9, Eagle Station, Indarch) all concentrations are relative to the total sample taken for the analysis. † Data from ref. 18. ‡ Duplicate chromium isotopic analysis. § 30% uncertainty (2σ). || 17% uncertainty (2σ). ¶ ρ is density. # Data from ref. 51.

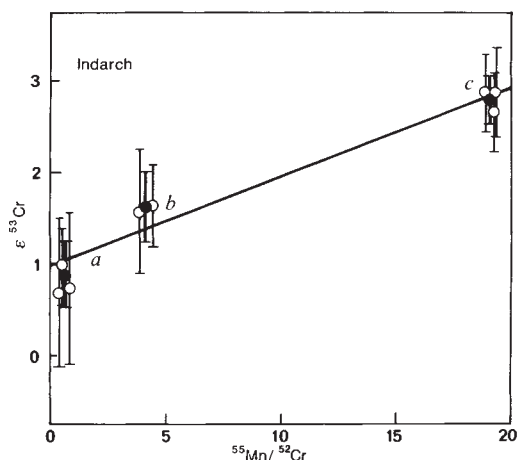


Fig. 1 ^{53}Mn - ^{53}Cr evolution diagram for the Indarch enstatite chondrite. The black dots are the weighted means of the individual measurements represented by white dots. The three groups of points are repeated measurements made on three different leachates: *a*, 37% HNO_3 ; *b*, 50% acetic acid; *c*, dilute acetic acid plus H_2O . The line drawn through the data has slope $^{53}\text{Mn}/^{55}\text{Mn} = (1.09 \pm 0.24) \times 10^{-6}$ and intercept $\epsilon^{53}\text{Cr} = 0.98 \pm 0.29$.

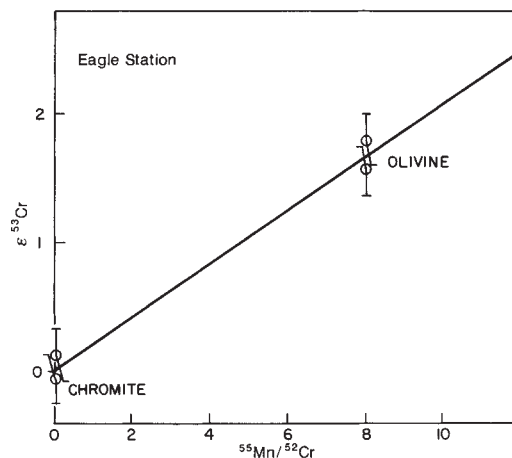


Fig. 2 ^{53}Mn - ^{53}Cr evolution diagram for the Eagle Station palisade. Only two minerals could be isolated in significant amounts. The line has slope $^{53}\text{Mn}/^{55}\text{Mn} = (2.3 \pm 0.3) \times 10^{-6}$ and intercept $\epsilon^{53}\text{Cr} = 0.03 \pm 0.14$.

Isochron or mixing line?

One question that arises is whether the correlation between $^{53}\text{Cr}/^{52}\text{Cr}$ and Mn/Cr is an isochron, or whether it simply reflects the mixing of two reservoirs of different nucleosynthetic history. A similar problem is encountered for the interpretation of the ^{26}Al - ^{26}Mg data obtained in refractory inclusions in chondrites^{3,21,25,26}. Although the experimental evidence for *in situ* decay is strong, this issue is still a subject of debate for ^{26}Al - ^{26}Mg ^{6,7,27}. We develop here a number of arguments in favour of an *in situ* decay of ^{53}Mn . As the $^{54}\text{Cr}/^{52}\text{Cr}$ ratio is similar in almost every non-FUN Allende inclusion^{17,18,28}, it is likely that the $^{53}\text{Cr}/^{52}\text{Cr}$ ratio was homogeneous in the reservoir producing these inclusions before the decay of ^{53}Mn . Further evidence concerns the $^{53}\text{Cr}/^{52}\text{Cr}$ ratio: for the iron-group elements the major isotopic effects in non-FUN inclusions are found for the most neutron-rich isotope, the other isotopes being essentially 'normal', considering the order of magnitude of the effect on the neutron-rich isotope. As ^{53}Mn is not a neutron-rich product, if the spread in $^{53}\text{Cr}/^{52}\text{Cr}$ did not originate in the decay of ^{53}Mn , this would constitute an additional difficulty in explaining the Cr isotopic data, in connection with Ti and Ca anomalies, in inclusions.

In the Indarch enstatite chondrite, equilibration between the sulphides occurred at temperatures higher than 1,000 °C^{29,30}. The partial melting of the sulphide phases has also been put forward as a likely possibility. As chromium and manganese are located in the sulphide phases in this class of meteorites, partial melting would result in isotopic homogeneity for these two elements. Therefore, *in situ* decay of ^{53}Mn is required to produce the correlation observed in Fig. 1.

Previous research on Eagle Station^{31,32} suggests a magmatic origin, which implies isotopic homogeneity at the time of formation. This conclusion is in agreement with the identical, although slightly anomalous, $^{54}\text{Cr}/^{52}\text{Cr}$ ratios measured for both the olivine and chromite fractions. This ^{54}Cr anomaly is small when compared to Allende inclusions. Its relation to the unusual oxygen isotopic composition²³ is not known yet. For Eagle Station, in our opinion, the correlation of $^{53}\text{Cr}/^{52}\text{Cr}$ with Mn/Cr can only be due to *in situ* decay of ^{53}Mn .

Thus, all of the data support the *in situ* decay of ^{53}Mn in all of the objects discussed here. This observation leads us to

consider the ^{53}Mn - ^{53}Cr couple as similar to the couple I-Xe, and to consider it likely that ^{53}Mn was ubiquitous.

In contrast to ^{107}Pd , for which the Pd/Ag ratios in meteorites allowed its detection only in iron meteorites, and the ^{26}Al detected only in refractory inclusions in chondrites^{3,4,21,25,26}, ^{53}Mn has also been detected in other meteorite classes. This strongly suggests that ^{53}Mn was widespread in the early Solar System. Thus, if these isotopes were from an exotic source, the injection of the latter might have affected a significant part of the Solar system. How homogeneously it was distributed in the Solar System is not yet well constrained.

Comparative systematics and chronology

To compare the results obtained on the various objects, we can use the slope and intercept values derived from the Mn-Cr correlation diagrams. The simplest case that can be addressed is that all the planetary objects derive from a unique, isotopically homogeneous reservoir evolving through time: one can hence write

$$\left(\frac{^{53}\text{Cr}}{^{52}\text{Cr}}\right)_\tau = \left(\frac{^{53}\text{Cr}}{^{52}\text{Cr}}\right)_* + \frac{^{55}\text{Mn}}{^{52}\text{Cr}} \left[\left(\frac{^{53}\text{Mn}}{^{55}\text{Mn}}\right)_* - \left(\frac{^{53}\text{Mn}}{^{55}\text{Mn}}\right)_\tau \right] \quad (2)$$

where the asterisk refers to the time of formation of the reservoir and τ to the time of formation of the meteorite. Using equation (2), the data points will fall on a straight line on the $(^{53}\text{Mn}/^{55}\text{Mn})$ - $(^{53}\text{Cr}/^{52}\text{Cr})$ diagram (Fig. 3) if the objects were produced from the same reservoir. In this diagram, $(^{53}\text{Mn}/^{55}\text{Mn})_\tau = (^{53}\text{Mn}/^{55}\text{Mn})_* e^{-\lambda \Delta \tau}$, where $\Delta \tau$ is the time interval between the formation of the reservoir and the formation of the planetary object, and λ is the decay constant of ^{53}Mn , $1.87 \times 10^{-7} \text{ yr}^{-1}$.

Although the value of $(^{53}\text{Mn}/^{55}\text{Mn})_*$ is unknown, we can calculate the age difference between two objects from:

$$\Delta \tau_{1-2} = \frac{1}{\lambda} \ln \left[\frac{(^{53}\text{Mn}/^{55}\text{Mn})_2}{(^{53}\text{Mn}/^{55}\text{Mn})_1} \right] \quad (3)$$

The $^{53}\text{Mn}/^{55}\text{Mn}$ axis in Fig. 3 can thus be used as a timescale.

We now consider the timescales involved for these samples. By considering the value of 4.4×10^{-5} for total rock plus inclusion BR1¹⁸ as the initial reference value, one can calculate ages relative to Allende. The Eagle Station meteorite yields an age 15.6 Myr younger, and Indarch is 19.6 Myr younger. These ages are of the same order as the difference between the age of formation of differentiated planetary objects, 4.555 Gyr³³, and that of the oldest Allende inclusions, 4.569 Gyr, measured by U-Pb dating³⁴. But these relative ages have to be considered

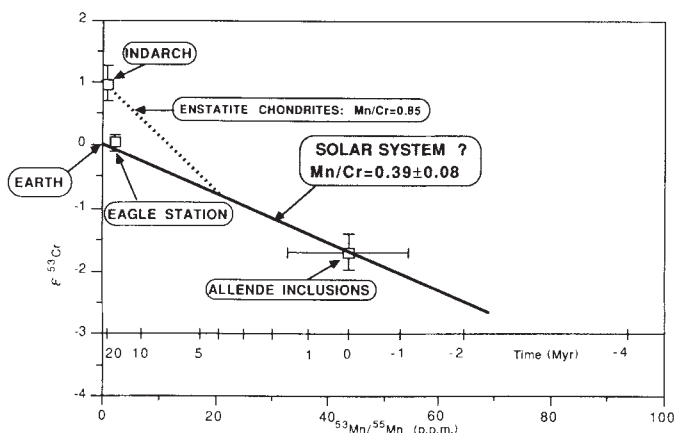


Fig. 3 Isotopic evolution of the source reservoir of the early Solar System material analysed here, assuming that Eagle Station, Indarch, the Earth and Allende were all derived from a single reservoir of Mn and Cr. The time origin is arbitrarily set at $^{53}\text{Mn}/^{55}\text{Mn} = 44$ p.p.m. and $\epsilon^{53}\text{Cr} = -1.7$.

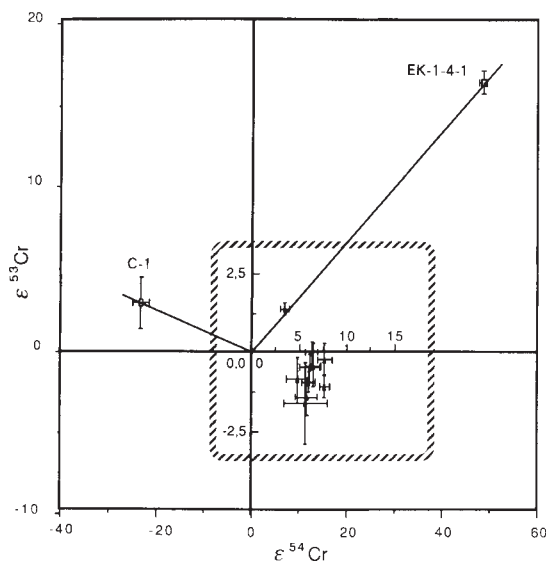


Fig. 4 $\epsilon^{53}\text{Cr}$ versus $\epsilon^{54}\text{Cr}$ in Allende inclusions. Mineral data are represented without their error bars. Data from ref. 28 for inclusions C-1 and EK-1-4-1 are also included for completeness.

with some caution. The initial assumption of the above calculation is that all objects investigated here are derived from the same isotopically homogeneous ^{53}Mn - ^{53}Cr reservoir. This is not necessarily true, as widespread heterogeneity has been demonstrated for a number of isotopes of the iron group: ^{48}Ca , ^{54}Cr and ^{50}Ti . Nevertheless, the same argument as for the *in situ* decay of ^{53}Mn in the inclusions can be used here: ^{53}Cr and ^{53}Mn are not neutron-rich isotopes; thus, except for the FUN inclusions, non-radiogenic nucleosynthetic effects should be minor or absent, thereby reinforcing the hypothesis of a solar nebula homogeneous in $^{53}\text{Cr}/^{52}\text{Cr}$ and $^{53}\text{Mn}/^{55}\text{Mn}$. FUN inclusions²⁸ in this case would be different from other Solar System materials in either their history or their origin, as they display ^{53}Cr effects which may not be related to ^{53}Mn decay. One possibility is that these excesses may be related to the not yet fully understood nucleosynthetic processes responsible for Ti anomalies: isotopic effects in ^{47}Ti and ^{49}Ti are of comparable magnitude to the effects in the neutron-rich ^{50}Ti . The conflict between the ^{53}Mn - ^{53}Cr chronology and the ^{129}I - ^{129}Xe

chronology³⁵ may not be a problem, as the physico-chemical properties of the two systems are very different.

The slope of the evolution line in Fig. 3 gives the Mn/Cr ratio of the Solar System. The value Mn/Cr = 0.39 is in agreement with the solar photospheric value³⁶, but significantly lower than the value obtained on CI chondrites³⁷. This may constitute a problem for the model of homogeneous $^{53}\text{Cr}/^{52}\text{Cr}$ and $^{53}\text{Mn}/^{55}\text{Mn}$ in the solar nebula. To explain the difference, Lee³⁸ has suggested a model of heterogeneous $^{53}\text{Cr}/^{52}\text{Cr}$ in the Solar System, with ^{53}Cr excesses relative to solar in Allende inclusions. This problem is not yet well constrained and we assume here that the Solar System was homogeneous for $^{53}\text{Mn}/^{55}\text{Mn}$.

As can be seen from Fig. 3, the Indarch enstatite chondrite has an initial $^{53}\text{Cr}/^{52}\text{Cr}$ ratio significantly higher than the solar value. This can be understood only if the parent material of this meteorite had a Mn/Cr ratio significantly higher than typical nebula values. It is well known that enstatite chondrites are depleted in refractory elements. If the source had a Mn/Cr ratio close to the enstatite chondrite value of ~ 1.0 , then this source had to be separated from the reservoir of solar composition only a few million years (~ 3 Myr) after the formation of Allende inclusions. This is still rather speculative, and further data are needed to document the evolution of the parent body or bodies of the enstatite chondrites. Nevertheless, the short time relative to Allende inclusions is in good agreement with strontium isotopic data³⁹, although the strontium ages based on initial ratios are model-dependent. The initial strontium isotope composition of enstatite chondrites is not resolved from that of Allende inclusions, implying that the two parent bodies were separated from the solar nebula within a few million years of each other.

The Eagle Station material as a whole has a high Mn/Cr ratio of ~ 8 , for which the $^{53}\text{Cr}/^{52}\text{Cr}$ ratio increases rapidly with time. If more than 2 Myr had elapsed between differentiation and closure of the Mn-Cr system, an excess relative to the solar evolution line (Fig. 3) would have been detected with the present experimental technique.

More complicated models than a single reservoir could be considered for the ^{53}Mn - ^{53}Cr evolution of the early Solar System. For example, mixing between reservoirs of different Mn and Cr isotope compositions would reduce the time differences given above by a factor which depends on the mixing ratio between the two reservoirs. For the present, however, the hypothesis that the solar nebula is a single ^{53}Mn - ^{53}Cr reservoir with homogeneous $^{53}\text{Cr}/^{52}\text{Cr}$ and $^{53}\text{Mn}/^{55}\text{Mn}$ is sufficient to explain the data.

In summary, less than 20 Myr elapsed between the formation of the Allende inclusions and the Indarch enstatite chondrite. Eagle Station, which probably formed at the core-mantle boundary of a small planetary body, was formed ~ 16 Myr after the Allende inclusions. This is a time range similar to what is obtained with ^{107}Pd - ^{107}Ag systematics in iron meteorites⁴⁰.

Comparison with iron-group isotope anomalies

One of the most interesting aspects of the ^{53}Mn - ^{53}Cr study is to try to relate this time-dependent system to the other isotope anomalies found in the iron group of nuclei. From nucleosynthetic arguments a certain coherence is expected in the anomalies⁸⁻¹¹.

Cr and Ti isotope compositions have been measured for the same Allende inclusions in which we measured the ^{53}Cr anomalies (Tables 1 and 2). The Cr and Ti concentrations in these inclusions were also determined using isotope dilution. This was done to examine whether any effect can be related to the concentration of the anomalous component(s).

There is no correlation between anomalies in ^{53}Cr and in ^{54}Cr (Fig. 4). If the source of variation in $^{53}\text{Cr}/^{52}\text{Cr}$ is the decay of ^{53}Mn , there is indeed no reason to expect a correlation between ^{53}Cr and ^{54}Cr . Although as mentioned above, some nucleosynthetic effects may be present for ^{53}Cr in some FUN inclusions²⁸,

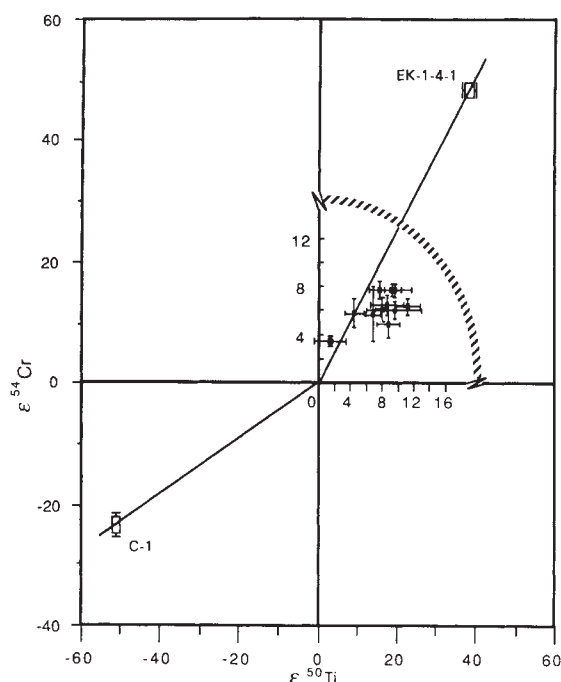


Fig. 5 Chromium versus titanium anomalies for the most neutron-rich isotopes of the two elements. Note that inclusion G6 has normal Ti isotope composition within error. Data from refs 15 and 28 (see Table 2) are included to cover the full range of variation.

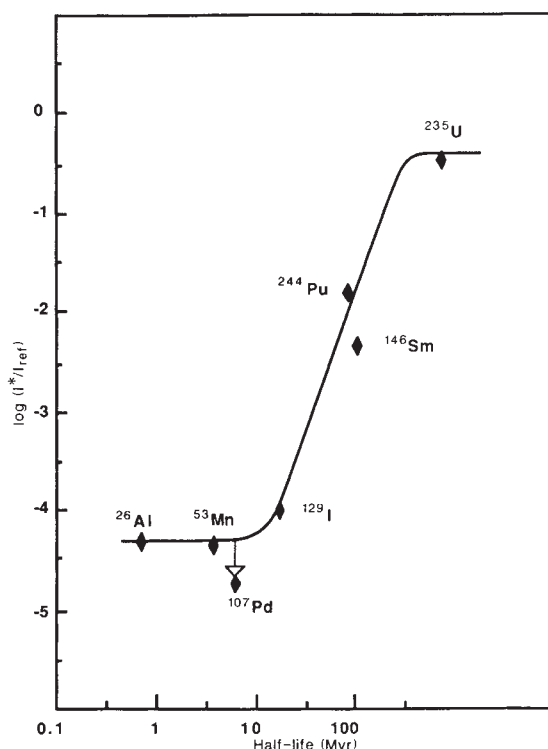


Fig. 6 Amounts of short-lived nuclides found in early Solar System materials, relative to a neighbouring major isotope of the same element (^{27}Al , ^{55}Mn , ^{108}Pd , ^{127}I and ^{144}Sm), except for ^{244}Pu and ^{235}U , which are referred to ^{238}U . For each nuclide, the abundance is plotted against the half-life.

these are always associated with non-neutron-rich effects in Ti. There is no indication of such an effect in our Ti data (Table 2).

In addition to the Allende inclusions, small ^{54}Cr excesses (Table 1) are also present in Indarch, Eagle Station, Murchison and the whole rock of Allende. The HF-insoluble residue of an Allende whole rock also displays this excess. This is not surprising, as this residue represents at least two-thirds of the Cr content of the whole rock. The fact that these meteorite samples have ^{54}Cr excesses that are smaller than those of Allende, but still clearly resolvable, indicates that there has been incomplete mixing of the originally heterogeneous solar nebula. This is also in agreement with widespread ^{50}Ti anomalies in bulk carbonaceous chondrites⁴¹.

The titanium isotope composition of the inclusions reported here shows ^{50}Ti excesses (Table 2) for all samples but one (Fig. 5). There is only a rough correlation between nonlinear effects in ^{50}Ti and ^{54}Cr (ref. 28). The orders of magnitude are similar but the correlation does not exist in detail. The remarkable feature is that one inclusion, G6, although having a ^{54}Cr excess, does not have a significant ^{50}Ti anomaly. This exception to the general rule^{12,13,15} can be explained by local inhomogeneities in the anomalous component, or by local chemical fractionation between Ti and Cr between the production of ^{50}Ti and ^{54}Cr near

the core of massive stars and their incorporation into the inclusion.

The isotope anomalies do not correlate with the concentrations of the corresponding element, nor can they be explained by the addition of pure ^{54}Cr and ^{50}Ti to matter with solar Ti and Cr isotope and chemical composition, in the case of non-FUN inclusions. This is not surprising, as even small variable chemical fractionation would destroy such correlations between different elements. Similar observations have been made for the effects in ^{48}Ca and ^{50}Ti ^{14,42}; for these isotopes there is always a rough correlation in magnitude and sign of the nonlinear effects, but the correlation fails in detail.

Several implications follow from the data:

(1) The ^{54}Cr and ^{50}Ti isotope anomalies are of the same order of magnitude. These nuclei most probably originate from a neutron-rich statistical equilibrium^{9-11,17} near the core of massive stars just before a supernova explosion⁸. For the FUN inclusions, additional nucleosynthetic processes have to be involved to produce the effects observed for the other isotopes¹⁵ of titanium. In the case of chromium²⁸ the problem is not yet solved, owing to the possibility of the decay of ^{53}Mn to ^{53}Cr ; the two remaining isotopes ^{50}Cr and ^{52}Cr are used for normalization.

(2) No Cr-Cr or Cr-Ti correlation exists in our data. This is in agreement with the fact that Allende inclusions do not result from a simple homogeneous binary mixture between two components, one 'normal' and the other 'abnormal'. The interpretation of Ti isotope data in Allende inclusions alone requires several components^{13,15,42}.

(3) Even if no detailed correlation can be established within the iron-group elements, there is always a rough correlation in sign and amplitude for the neutron-rich isotopes ^{48}Ca , ^{50}Ti and ^{54}Cr ^{28,42}.

In view of the above, the hypothesis of injection of an abnormal component and mixture in a gaseous state, followed by a

Table 2 Titanium isotopic data

Sample	Concentration (p.p.m.)	$\epsilon^{47}\text{Ti}$	$\epsilon^{49}\text{Ti}$	$\epsilon^{50}\text{Ti}$
G1 mean	5,580	0.2 ± 0.8	-0.2 ± 0.7	9.3 ± 1.1
G2 mean	9,133	-1.1 ± 1.2	-0.7 ± 1.4	10.3 ± 1.6
G6 mean	8,573	0.4 ± 1.2	-1.7 ± 1.7	0.5 ± 1.7
G4	7,255	0.3 ± 0.6	0.0 ± 0.7	8.0 ± 0.8
SB	7,475	0.7 ± 1.2	0.4 ± 1.6	11.2 ± 1.6
BR1 mean	7,231	-0.6 ± 0.9	-1.2 ± 1.0	8.7 ± 1.2
BR9	6,831	-1.0 ± 1.6	0.4 ± 1.9	9.8 ± 2.1

Titanium isotopic data are normalized to $^{46}\text{Ti}/^{48}\text{Ti} = 0.108548$ (ref. 15). The results are quoted in ϵ units relative to terrestrial laboratory standards.

series of condensations of solid minerals, has to be discarded, as it would result in simple correlations. On the other hand, the results obtained seem to indicate that nucleosynthesis was followed immediately by a chemical fractionation of these elements, probably as condensed micrograins⁴³. The latter would have then been mechanically mixed in the refractory inclusions or followed by more complex processes such as differentiation in objects such as Indarch or Eagle Station.

Origin of the short-lived nuclides

Our results do not seem to be consistent with a fossil origin of the short-lived nuclides as proposed by Clayton²⁷. The reasonable slope of the ⁵³Mn–⁵³Cr isochrons found in objects produced in various physico-chemical contexts argue strongly against this possibility. As *in situ* decay of ⁵³Mn has been clearly identified in evolved objects of the Solar System, this seems the most straightforward interpretation of the Allende inclusion data as well. Although the relation with precise lead chronology³⁴ is not yet clear, Allende inclusions display ages within 20 Myr of other primitive Solar System material.

We now plot the abundance of each short-lived nuclide^{6,7,18}, relative to a major stable isotope of the same element produced in similar amounts at their stellar production sites, versus the half-life of the short-lived nuclide (Fig. 6). A regular pattern is observed, with a discontinuity starting at ¹²⁹I. For the shorter half-lives, ²⁶Al, ¹⁰⁷Pd and ⁵³Mn have comparable abundances. Extensive reviews about the behaviour of short-lived nuclides have already been published^{6,7}. According to nucleosynthesis calculations^{8,45,46}, the production ratios of ²⁶Al, ⁵³Mn, ¹⁰⁷Pd and ¹²⁹I relative to their major stable neighbours ²⁷Al, ⁵⁵Mn, ¹⁰⁸Pd and ¹²⁷I, respectively, are within the same order of magnitude. Thus, the flat pattern (Fig. 6) between ²⁶Al and ¹²⁹I could be explained by the addition of about 10⁻³ to 10⁻⁴ solar masses of freshly synthesized matter ~1–3 Myr before the formation of the small planetary objects. ¹⁰⁷Pd plots below ²⁶Al and ⁵³Mn; in contrast to the latter isotopes, ¹⁰⁷Pd has been found not in Allende inclusions but in iron meteorites, which are differentiated objects; thus there may have been significant decay of ¹⁰⁷Pd^{5,40} before closure of the Pd–Ag system. A delay of a few million years is needed to allow ⁴¹Ca to decay to an undetectable level⁴⁷. The last substantial addition or the end of addition of true r-process heavy nuclei such as ²⁴⁴Pu took place a few hundred million years before the formation of the Solar System. From the shape of the curve, this contribution to r-process nuclei

would represent a few per cent^{6,7} of the same nuclei already in the nebula. At the low end of the timescale, a time delay of between 10⁵ and ~10⁶ yr is needed between nucleosynthesis and the formation of planetary objects⁴⁴. This is in good agreement with the complete decay of ⁴¹Ca (ref. 47), but if the smallest model times are correct, ⁴¹Ca could still be present in Allende inclusions. It must be kept in mind that ⁴¹Ca decays into ⁴¹K, a rather volatile nuclide; on the other hand, it is known from oxygen data⁴⁸ that exchanges with normal Solar System matter have occurred at intermediate temperatures. Those exchanges may well have also erased ⁴¹K isotope anomalies without affecting the more refractory Al–Mg and Mn–Cr systems.

It should be kept in mind, however, that silver is even more volatile⁴⁹ than potassium, and if potassium isotope anomalies have been erased by isotopic rehomogenization in nebular processes, this is likely to happen to silver as well. This is probably another major problem for the 'cosmic chemical memory' model²⁷ in the interpretation of the isotope effects related to the decay of short-lived nuclides other than ²⁶Al. In our opinion, the support of the experimental data for the model described here is, together with petrographic data, an argument in favour of *in situ* decay of ²⁶Al in the Allende inclusions. Nevertheless, there are a number of features which need further investigation: the nucleosynthetic sources of the different short-lived nuclides are not necessarily related; and ²⁶Al is mostly synthesized in novae⁴⁵, whereas ⁵³Mn is more probably a supernova product^{8,9}.

It is an open question whether several nucleosynthetic episodes of different nature contributed to the solar nebula's short-lived nuclides in the last ten million years before the formation of small planetary bodies. The possibility exists that the pattern obtained in Fig. 6 is a fortuitous mixture of matter from several stellar nucleosynthetic sources. The key feature is that all of the physico-chemical processes between nucleosynthesis and formation of small planetary bodies had to happen within a few tens of million years for ¹⁰⁷Pd and ⁵³Mn, and a few million years for ²⁶Al.

Samples were provided by E. A. King, P. Pellas and J. Halbout. We thank D. Laverne and G. Michard for their Mn concentration measurements; G. J. Wasserburg, D. Papanastassiou, G. Lugmair, H. Wanke, D. D. Clayton and F. Begemann for interesting discussions; W. G. Pegram and K. V. Ragnarsdottir for improving the English; C. Mercier for typing the manuscript and the technical staff of the laboratory for the help provided during this study.

Received 17 September; accepted 31 December 1987.

- Clayton, R. N., Grossman, L. & Mayeda, T. K. *Science* **182**, 485–488 (1973).
- Begemann, F. *Rep. Prog. Phys.* **43**, 1309–1356 (1980).
- Gray, C. M. & Compston, W. *Nature* **251**, 495–497 (1974).
- Lee, T. & Papanastassiou, D. A. *Geophys. Res. Lett.* **1**, 225–228 (1974).
- Kelly, W. R. & Wasserburg, G. J. *Geophys. Res. Lett.* **5**, 1079–1082 (1978).
- Wasserburg, G. J. & Papanastassiou, D. A. in *Nuclear Astrophysics* (eds Barnes, C. A., Clayton, D. D. & Shramm, D. N.) 77–140 (Cambridge University Press, 1982).
- Wasserburg, G. J. in *Protostars and Planets II* (eds Black, D. C. & Matthews, M. S.) 703–737 (University of Arizona Press, Tucson, 1985).
- Woosley, S. E., Arnett, W. D. & Clayton, D. D. *Astrophys. J. Suppl. Ser.* **26**, 231–312 (1973).
- Hainebach, K. L., Clayton, D. D., Arnett, W. D. & Woosley, S. E. *Astrophys. J.* **193**, 157–168 (1974).
- Hartmann, P., Woosley, S. E. & El Eid, M. F. *Astrophys. J.* **297**, 837–845 (1985).
- Cameron, A. G. W. *Astrophys. J.* **230**, L53–L57 (1979).
- Heydegger, H. R., Foster, J. J. & Compston, W. *Nature* **278**, 704–707 (1979).
- Niemeyer, S. & Lugmair, G. W. *Geochim. cosmochim. Acta* **48**, 1401–1416 (1984).
- Junck, M. H. A., Shimamura, T. & Lugmair, G. W. *Geochim. cosmochim. Acta* **48**, 2651–2658 (1984).
- Niederer, F. R., Papanastassiou, D. A. & Wasserburg, G. J. *Geochim. cosmochim. Acta* **49**, 835–851 (1985).
- Lee, T., Papanastassiou, D. A. & Wasserburg, G. J. *Astrophys. J.* **220**, L21–L25 (1978).
- Birck, J. L. & Allègre, C. J. *Geophys. Res. Lett.* **11**, 943–946 (1984).
- Birck, J. K. & Allègre, C. J. *Geophys. Res. Lett.* **12**, 745–748 (1985).
- Birck, J. L. in *Advances in Nuclear Astrophysics* (eds Vangioni-Flam, F., Audouze, J., Casse, M., Chieze, J. P. & Tram Thau Van, J. P.) 461–469 (Frontière, Paris, 1986).
- Reynolds, J. H. *Phys. Rev. Lett.* **4**, 8–10 (1960).
- Lee, T., Papanastassiou, D. A. & Wasserburg, G. J. *Geophys. Res. Lett.* **3**, 109–112 (1976).
- Robert, F., Merlivat, L. & Javoy, M. *Nature* **282**, 785–789 (1979).
- Clayton, R. N. & Mayeda, J. K. *Earth planet. Sci. Lett.* **40**, 168–174 (1978).
- Brigham, C. A., Hutcheon, I. D., Papanastassiou, D. A. & Wasserburg, G. J. *Lunar planet. Sci.* **18**, 85–86 (1986).
- Stegmann, W. & Begemann, F. *Earth planet. Sci. Lett.* **55**, 266–272 (1981).
- Hinton, R. W. & Bischoff, A. *Nature* **308**, 169–172 (1984).
- Clayton, D. D. *Astrophys. J.* **310**, 490–498 (1986).
- Papanastassiou, D. A. *Astrophys. J.* **308**, L27–L30 (1986).
- Mason, B. *Geochim. cosmochim. Acta* **30**, 23–39 (1966).
- Hohenberg, C. M., Podosek, F. A. & Reynolds, J. H. *Science* **156**, 233–236 (1967).
- Davis, A. M. thesis, Yale Univ. (1977).
- Scott, E. R. D. *Geochim. cosmochim. Acta* **41**, 693–710 (1977).
- Chen, J. H. & Wasserburg, G. J. *Lunar planet. Sci.* **16**, 119–120 (1985).
- Manhès, G., Göpel, C. & Allègre, C. J. (abstr.) *Terra cognita* **6**, 173 (1986).
- Lewis, R. S. & Anders, E. *Proc. natn. Acad. Sci. U.S.A.* **72**, 268–273 (1975).
- Ross, J. E. & Aller, L. H. *Science* **191**, 1223–1229 (1976).
- Anders, E. & Ebihara, M. *Geochim. cosmochim. Acta* **46**, 2362–2380 (1982).
- Lee, T. *Lunar planet. Sci.* **18**, 548–549 (1987).
- Minster, J. F., Birck, J. L. & Allègre, C. J. *Nature* **300**, 414–418 (1982).
- Chen, J. H. & Wasserburg, G. J. *Geochim. cosmochim. Acta* **47**, 1725–1737 (1983).
- Niemeyer, S. *Geophys. Res. Lett.* **12**, 733–736 (1985).
- Zinner, E. K., Fahey, A. J., Goswami, J. N., Ireland, T. R. & McKeegan, K. J. *Astrophys. J.* **311**, L103–L107 (1986).
- Liffman, K. & Clayton, D. D. *Lunar planet. Sci.* **18**, 522–553 (1987).
- Cameron, A. G. W. in *Protostars and Planets II* (eds Black, D. C. & Matthews, M. S.) 1073–1099 (University of Arizona Press, Tucson, 1985).
- Clayton, D. D. *Astrophys. J.* **280**, 144–149 (1984).
- Clayton, D. D. & Fowler, W. A. *Ann. Phys.* **16**, 51–68 (1961).
- Hutcheon, I. E., Armstrong, J. T. & Wasserburg, G. J. *Meteoritics* **19**, 243–244 (1984).
- Clayton, R. N. & Mayeda, T. K. *Geophys. Res. Lett.* **4**, 295–298 (1977).
- Larimer, J. W. *Geochim. cosmochim. Acta* **31**, 1215–1238 (1967).
- Shields, W. R., Murphy, J. J., Cantazaro, E. J. & Garner, L. J. *J. Res. natn. Bur. Stand.* **70A**, 193–197 (1966).
- Kallemeyn, G. W. & Wasson, J. T. *Geochim. cosmochim. Acta* **45**, 1217–1230 (1981).

Positron emission tomographic studies of the cortical anatomy of single-word processing

S. E. Petersen, P. T. Fox, M. I. Posner, M. Mintun & M. E. Raichle

Department of Neurology and Neurological Surgery, Washington University School of Medicine, 660 South Euclid Avenue, St Louis, Missouri 63110, USA

The use of positron emission tomography to measure regional changes in average blood flow during processing of individual auditory and visual words provides support for multiple, parallel routes between localized sensory-specific, phonological, articulatory and semantic-coding areas.

LANGUAGE is an essential characteristic of the human species, and has been studied by disciplines ranging from philosophy to neurology. Because language is so complex, cognitive and neurological studies often focus on processing of individual words (lexical items). Cognitive models for lexical processing consider words perceived visually and auditorily to involve separate modality-specific codes, with access in parallel to shared output (articulatory) and meaning (semantic) codes¹⁻⁶. In contrast, the model most widely accepted in the clinical neurological literature argues for serial processing, with an early recoding of visual input into an auditory-based code which is used in turn for semantic and articulatory access^{7,8}.

We have used recent advances in the precision of positron emission tomography (PET) for measuring activity-related changes in regional cerebral blood flow to identify brain regions active during three levels of single-word processing. Our results indicate localization of different codes in widely separated areas of the cerebral cortex. The results favour the idea of separate brain areas involved in separate visual and auditory coding of words, each with independent access to supramodal articulatory and semantic systems. These findings fit well with the parallel models, but argue against the obligatory visual-to-auditory recoding and serial nature of the clinical neurological models.

Methods

Brain blood flow was measured in 17 (11 female, 6 male) right-handed normal volunteers using a bolus intravenous injection of ¹⁵O-labelled water (half-life, 123 s) and a 40-s data acquisition^{9,10}. A series of 6–10 blood flow scans were obtained in each subject (10 m interscan interval). Within this series, conditions were designed as a hierarchy of paired comparisons to allow subtractive (task minus control) data analysis (see below).

Stimuli were presented throughout data acquisition. All stimuli were frequent English nouns presented at a rate of one

per second. Visually presented words appeared on a colour monitor suspended 300 mm from the subject. Auditory words were presented through hearing-aid type speakers fitted within the ears and driven by a digital tape recorder.

Four behavioural conditions formed a three-level subtractive hierarchy (Table 1). Each task state was intended to add a small number of operations to those of its subordinate (control) state¹¹. In the first-level comparison, the presentation of single words without a lexical task was compared to visual fixation without word presentation. Note that no motor output or volitional lexical processing was required in this task; rather, simple

Table 2 Sensory tasks

Region	Coordinates (mm)			Magnitude
	Z	X	Y	
Visual				
1. Striate cortex (L)	10	6	-72	2.28†
2. Striate cortex (R)	10	-12	-72	2.66†
3. Extrastriate cortex (L)	2	24	-58	3.82‡
4. Extrastriate cortex (R)	6	-26	-66	2.95‡
5. Inferior lateral occipital cortex (R)	-4	-34	-46	3.38‡
Auditory				
6. Posterior superior temporal cortex (L)	14	46	-10	2.46†
7. Temporal cortex (R)	12	-42	-16	2.76‡
8. Anterior superior temporal cortex (L)	-2	42	10	3.02†
9. Temporoparietal cortex (L)	14	54	-30	2.88‡
10. Lateral temporal cortex (R)	8	-62	-12	3.30‡
11. Inferior anterior cingulate cortex (L)	18	12	44	2.34†

Subtraction conditions: Passive words – Fixation point. For Tables 2–4, the following conventions are used: the region is given a mnemonic anatomical name associated with the coordinates. The coordinates and magnitudes of response are determined using a three-dimensional search algorithm on the averaged subtraction image. The coordinates are in mm from a 0, 0, 0 point that is at the level of a line drawn between the anterior and posterior commissures ($z = 0$), at the mid-line of the brain ($x = 0$), and located antero-posteriorly halfway between the commissures ($y = 0$). The magnitudes are the change in blood flow in ml/(100 g × min), and the statistical significance of the points is assessed with a two-stage testing procedure. The distribution of the magnitudes of local blood-flow change is tested for outliers using an omnibus gamma-2 test. For all averaged images presented here, there are statistically significant outliers. The foci with the largest magnitude of blood-flow change are then given a z-score with respect to the population of all local changes within an image. All foci of change with a P -value < 0.03 are reported in the tables. † $P < 0.03$, ‡ $P < 0.01$.

In general, the passive presentation subtractions identify modality-specific foci of activation, whereas the higher level subtractions activate similar regions across modalities.

Table 1 Paradigm design

Subtraction	Control state	Stimulated state	Task
Sensory task	Fixation point only	Passive words	Passive sensory processing Modality-specific word code
Output task	Passive words	Repeat words	Articulatory code Motor programming Motor output
Association Task	Repeat words	Generate uses	Semantic association Selection for action

The rationale of the three levels stepwise paradigm design is shown. At the second and third level, the control state is the stimulated state from the previous level. Some hypothesized cognitive operations are represented in the third column.

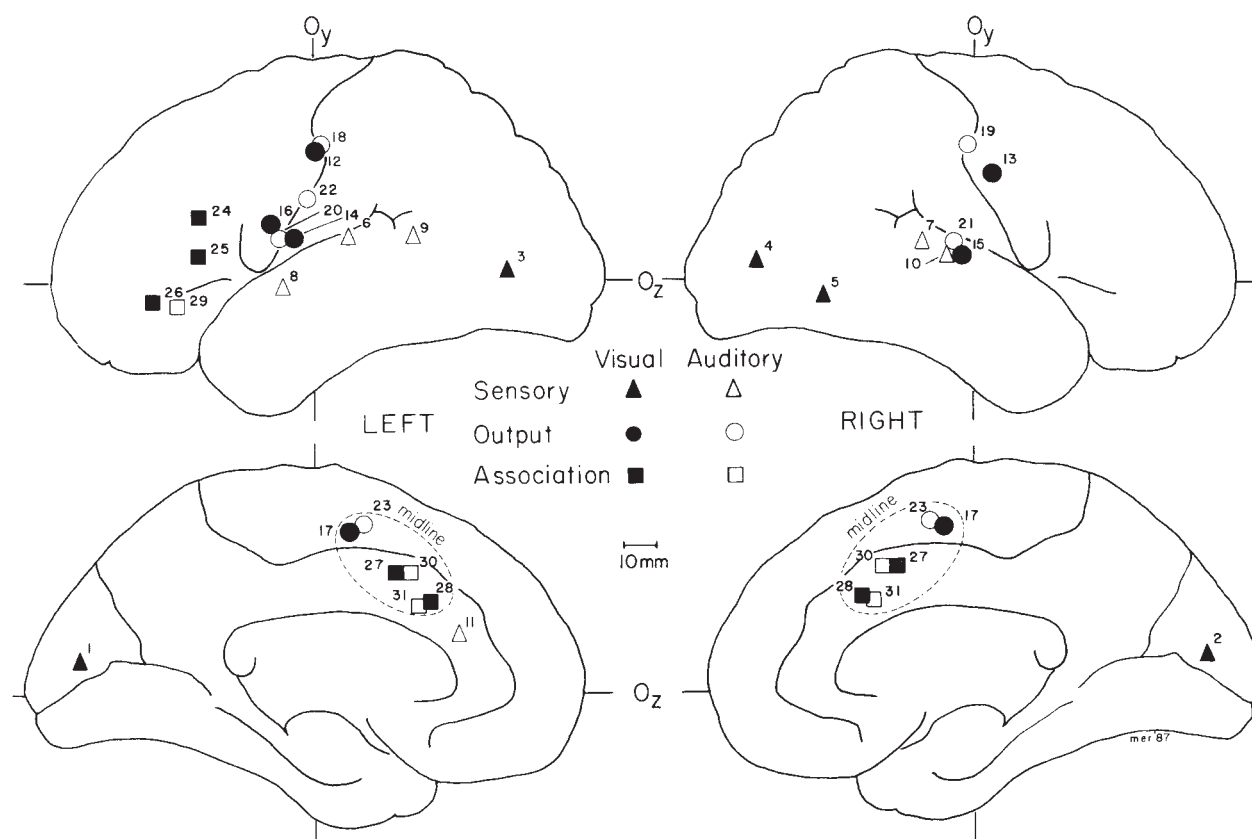


Fig. 1 Schematic lateral (upper) and medial (lower) surface views of the left and right hemispheres with superimposed cortical activation foci. O_y and O_z are 0 reference planes. Each numbered symbol represents a cortical focus of activation, the number referring to Focus number in Tables 2-4. The key to the activation conditions is in the figure. Notice that for passive subtractions, there is no overlap between visual (filled triangle) and auditory (open triangle) sensory tasks. There is considerable overlap, however, between presentation modality for the association and output foci.

sensory input and involuntary word-form processing were targeted by this subtraction (sensory task). In the second-level comparison, speaking each presented word was compared with word presentation without speech. Areas involved in output coding and motor control were targeted by this comparison. In the third-level comparison, saying a use for each presented word (for example, if 'cake' was presented, to say 'eat') was compared with speaking presented words. This comparison targeted areas involved in the task of semantic processing (verb-noun association), as distinct from speech, sensory input and involuntary word-form processing (association task).

Images were analysed by paired (intrasubject) subtraction. Task-state minus control-state subtractions created images of the regional blood flow changes associated with the operations of each cognitive level. Intersubject averaging was used to increase the signal-to-noise ratio of these subtracted images¹². Averaging required anatomical standardization of all images; this was based on a previously described stereotactic method of anatomical localization for PET images^{13,14}.

Statistical significance was determined by distribution analysis of the entire population (both positive and negative) of independent regional changes within each averaged subtracted image. The location and magnitude of these changes were determined using a centre-of-mass computer search algorithm¹⁵. Each change distribution contained both noise and task-induced responses. During averaging, task-induced responses gained in magnitude relative to image noise, becoming 'outliers' in the distribution¹². Significant responses, then, were defined using tests for outlier detection¹⁶. Statistical analysis was two-tiered: first, omnibus testing (gamma-2 statistic) determined whether an image (a distribution) contained any significant responses

(any outliers); then, post-hoc analysis by Z-score ascribed significance levels to each response within the population. All distributions reported had a gamma-2 significance level of $P < 0.05$. All cortical responses with a Z-score over 2.17 ($P < 0.03$) are reported.

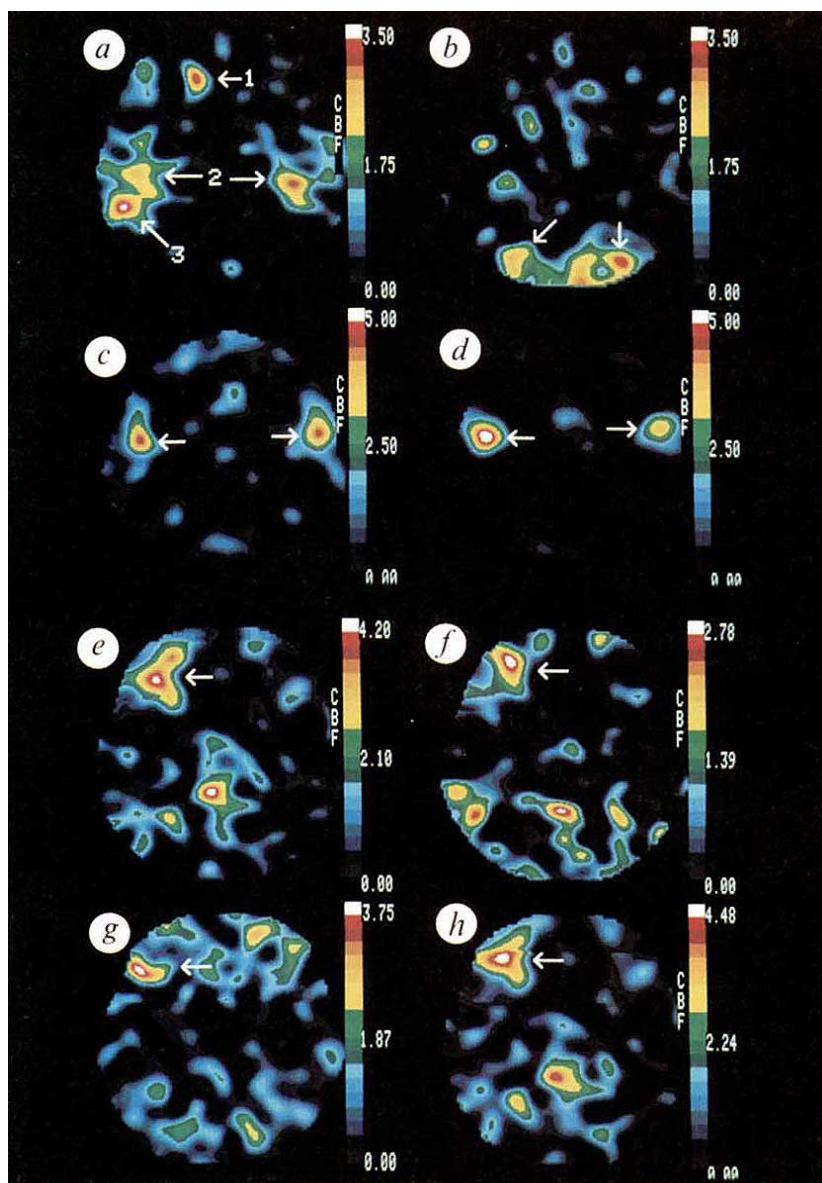
Lexical processing regions

Regions of activation are enumerated in Tables 2-4, and the cortical sites of activation are summarized in Fig. 1. The most striking aspect of Fig. 1 is that there are relatively few areas of activation added by each task and that these areas are clustered in a few critical parts of the cortex.

Modality-specific primary and non-primary sensory regions were activated by passive auditory or visual presentation of words (Table 2, Fig. 2a and b). No regions were activated for both auditory and visual presentation. The areas identified appear to support two different computational levels in each modality, one of passive sensory processing and a second level of modality-specific word-form processing.

For the visual modality, the main cortical activations are in the striate cortex and in a small set of prestriate areas reaching as far anterior as the temporal-occipital boundary. The primary striate responses were similar to those produced by other types of visual stimuli^{17,18}. However, the regions of extrastriate occipital cortex in Table 2 have so far been activated only by the presentation of visual words. These regions may represent a network which codes for visual word form. Lesions near these regions sometimes cause pure alexia, that is, the inability to read words without other language deficits^{19,20}. According to some cognitive models^{1,3,21}, a visual word form would be generated by a cooperative computational network including feature,

Fig. 2 *a* and *b*, Auditory versus visual comparison. A horizontal slice through averaged subtraction image represents blood-flow change when blood-flow during fixation is subtracted from blood flow present during presentation of word stimuli at 1 Hz (sensory task). Slice in *a* and *b* is taken 1.6 cm above AC-PC line. Foci of activity present at this level include temporoparietal cortex, bilateral superior posterior temporal cortex, inferior anterior cingulate for auditory presentation, and some occipital cortical activation for visual presentation. Note the non-overlapping distributions of activity for visual and auditory presentation in *a* and *b* during passive presentation. *c* and *d*, Auditory versus visual comparison. A horizontal slice through an averaged subtraction image representing blood-flow change when blood flow during passive presentation of words is subtracted from blood flow during vocal repetition of presented words (output task). Slice is taken 4.0 cm above AC-PC line. The foci present for both auditory and visual presentation are located on rolandic cortex, just anterior and superior to regions activated by somatosensory stimulation of the lips and probably represent the mouth representation of primary motor cortex. *e* and *f*, Auditory versus visual comparison. A horizontal slice through an averaged subtraction image representing blood-flow change when blood flow during repetition of presented words is subtracted from blood flow during vocalization of an appropriate use for the presented word (such as presentation of 'cake' ... output might be 'eat') (cognitive subtraction). Slice is taken 0.8 cm below AC-PC line. Foci for both presentation modalities occur in inferior anterior frontal cortex, probably area 47 of Brodmann. Those areas of activation are strongly left-lateralized. *g* and *h*, Comparison of activation in two semantic tasks. The slice on the right (*h*) is from the same condition as *e*; *g*, the blood-flow change when the blood flow during passive presentation of words at 2.5 Hz is subtracted from blood flow during a condition where the subject is asked to monitor this string of words for members of a specific semantic category. In the semantic monitoring task, there is no motor output during the scan. Subjects are asked after the scan for a gross estimate of the percentage of target words. The similar foci of activation in these two different semantic tasks implicate this region in semantic processing. Slice is taken 0.6 cm below AC-PC line.



letter, and word levels. The multiple areas activated could represent the different levels of such a network.

For auditory processing, areas of activity were found bilaterally in primary auditory cortex, and left-lateralized in temporoparietal cortex, anterior superior temporal cortex, and inferior anterior cingulate cortex. The temporoparietal and anterior superior temporal regions have not been activated by presentation of non-word auditory stimuli²²⁻²⁴. The temporoparietal region is near the angular and supramarginal gyri, areas that have been associated in lesion studies with the phonological deficits^{25,26}, and is a good candidate for a phonological coding region.

Areas related to motor output and articulatory coding are activated when words are repeated aloud (Table 3). In general, similar regions were activated for visual and auditory presentation. The activated regions included primary sensorimotor mouth cortex at a location corresponding to previous descriptions of sensorimotor topography²⁷. Also activated were a set of premotor structures including a midline structure (supplementary motor area, SMA) and a set of activations around the sylvian fissure. The left sylvian regions are near Broca's area, a region often viewed as specifically serving language output^{7,8}. But sylvian activation was also found in the right hemisphere,

and this bilateral sylvian activation was also found when subjects were instructed to simply move their mouths and tongues, arguing against specialization of this region for speech output. Small lesions confined to classically defined Broca's area most frequently cause stuttering and oral apraxia rather than full-blown Broca's aphasia²⁸, adding further support to the view that these regions are related to general motor, rather than language-specific output programming.

The association tasks activated two areas of cerebral cortex for both auditory and visual presentation. A left inferior frontal area was identified that almost certainly participates in processing for semantic association. The second area, anterior cingulate gyrus, appears to be part of an anterior attentional system engaged in selection for action. This localization of function was suggested by the performance of a converging experiment in which subjects monitored lists of words for members of a semantic category (such as monitoring for dangerous animals). In the semantic monitoring condition, left-frontal activation was strong and was unaffected by the number of targets in the list, supporting a semantic-processing function. Anterior-cingulate activation, however, was much stronger for lists containing many targets than for those with few, suggesting activation only when target selection was frequent. Similarly, rapid uncued move-

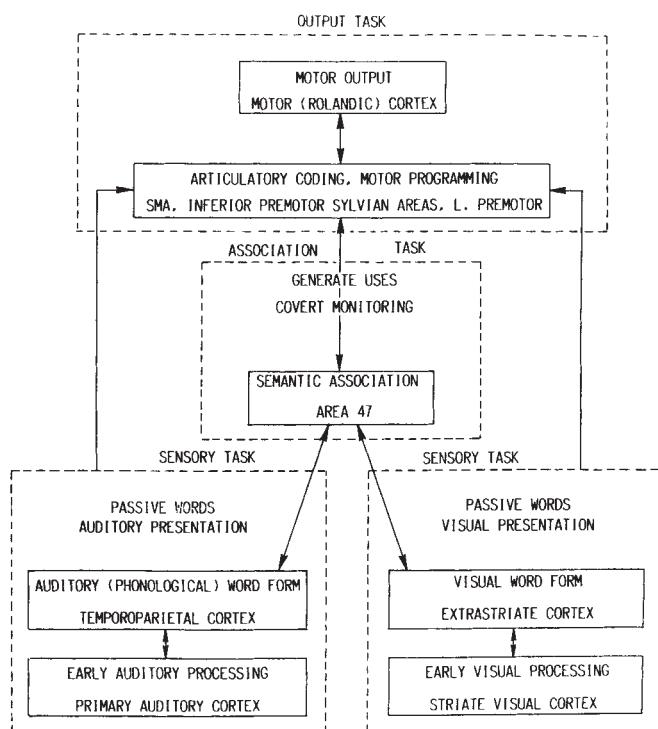


Fig. 3 A general network relating some of the areas of activation in this study to the different levels of lexical processing. There are many alternative networks consistent with the conditions under which the areas are activated, but this arrangement represents a simple design consistent with our results, and some convergent experiments from other types of studies. The dashed boxes outline the different subtractions. The solid boxes outline possible levels of coding and associated anatomical areas of activation.

ments and imagined movements activated anterior cingulate²⁹, whereas monitoring very low-frequency non-linguistic visual stimuli (J. V. Pardo, P.T.F., M.E.R., unpublished observations) activated neither cingulate nor left-frontal cortex. In accord with these observations, lesions of the anterior cingulate reduced the frequency of movements and speech (akinetic mutism)³⁰⁻³², whereas left-frontal lesions produced deficits in word-fluency tests³³, and in semantic-priming tasks^{34,35}.

Lexical processing models

What type of model do these results support? A serial single-route model has been widely accepted in clinical neurology^{7,8}. In the serial model, access to semantics is by a phonological code, and access to output is by semantics. Thus, a visual word must be phonologically recoded (said to occur in the angular gyrus) and must establish semantic associations (Wernicke's area in the posterior temporal lobe) before output coding. Our results are more consistent with multiple-route models in concept^{1-6,36-38}, and are also quite inconsistent with the serial neurological model in detail.

First, there is no activation in any of our visual tasks near Wernicke's area or the angular gyrus in posterior temporal cortex. Visual information from occipital cortex appears to have access to output coding without undergoing phonological recoding in posterior temporal cortex. Second, tasks calling for semantic processing of single words activate frontal, rather than posterior, temporal regions. Third, sensory-specific information appears to have independent access to semantic codes and output codes; simple repetition (output tasks) of a presented word failed to activate the left-frontal semantic area (association tasks). A framework consistent with these results is presented in Fig. 3.

Table 3 Output tasks

Region	Coordinates (mm)			Magnitude
	Z	X	Y	
Visual				
12. Mouth region, rolandic cortex (L)	40	46	0	4.34‡
13. Rolandic cortex (R)	32	−52	6	3.46‡
14. Buried sylvian cortex (L)	14	31	6	3.04†
15. Lateral sylvian cortex (R)	8	−63	−4	2.96†
16. Premotor cortex (L)	18	48	14	2.98†
17. Supplementary motor area (SMA)	50	−2	10	3.36†
Auditory				
18. Mouth region, rolandic cortex (L)	42	46	−2	3.64‡
19. Rolandic cortex (R)	40	−56	2	3.78‡
20. Buried sylvian cortex (L)	14	34	10	3.17†
21. Lateral sylvian cortex (R)	12	−62	−7	3.22‡
22. Premotor cortex (L)	26	52	2	3.06†
23. SMA	52	2	14	2.80†

Subtraction conditions: Repeat words - Passive visual words. See Table 2 legend for details of conventions used.

Table 4 Association tasks

Region	Coordinates (mm)			Magnitude
	Z	X	Y	
Visual				
24. Dorsolateral pre-frontal cortex (L)	20	44	36	2.98‡
25. Lateral prefrontal cortex (L)	8	38	36	2.96‡
26. Inferior prefrontal Cortex (L)	−6	−28	50	2.26†
27. Anterior cingulate	38	−6	24	3.12‡
28. Inferior anterior cingulate	28	−2	34	2.76‡
Auditory				
29. Inferior prefrontal cortex (L)	−6	33	43	3.10‡
30. Anterior cingulate	38	7	28	3.28‡
31. Inferior anterior cingulate	28	11	31	3.04‡

Subtraction conditions: Generate words - Repeat visual words. See Table 2 legend for details of conventions used.

The combination of cognitive and neurobiological approaches, of which this study is an example, has given us information about the functional anatomy of perception, attention, motor control, and language. As these endeavours proceed, solutions to the problem of mind-brain interaction that have intrigued us for so long should be illuminated.

This work was supported in part by the Office of Naval Research the National Institute of Health, and by the McDonnell Center for Higher Brain Function and the MacArthur Foundation.

Received 6 October 1987; accepted 5 January 1988.

1. LaBerge, D. & Samuels, J. *Cognitive Psychol.* **6**, 293-323 (1974).
2. Rumelhart, D. E. & McClelland, J. L. *Psychol. Rev.* **89**, 60-94 (1982).
3. Rumelhart, D. E. & McClelland, J. L. *Parallel Distributed Processing Vols 1 and 2* (MIT, Cambridge, 1986).
4. Carr, T. H. & Pollatsek, A. *Reading Research Vol. 5*, 1-82 (Academic, New York, 1985).
5. Posner, M. I. in *Chronometric Explorations of Mind* (Posner, M. I. & Marin, O. S. M.) (Erlbaum, Englewood Height, N.J., 1978).
6. Coltheart, M. *Attention and Performance XI* (Erlbaum, Hillsdale, N.J., 1985).

7. Geschwind, N. *Brain* **88**, 237–294, 585–644 (1965).
8. Geschwind, N. *Scient. Am.* **241**, 158–168 (1979).
9. Herscovitch, P., Markham, J. & Raichle, M. E. *J. Nucl. Med.* **24**, 782–789 (1983).
10. Raichle, M. E., Martin, W. R. W., Herscovitch, P., Mintun, M. A. & Markham, J. *J. Nucl. Med.* **24**, 790–798 (1983).
11. Sternberg, S. *Acta Psychol.* **30**, 276–315 (1969).
12. Fox, P. T., Mintun, M. A. & Raichle, M. E. *J. cerebral Blood Flow Metab.* (in the press).
13. Fox, P. T., Perlmutter, J. S. & Raichle, M. E. *J. Comp. Assist. Tomogr.* **9**, 141–153 (1985).
14. Tailarach, J. *et al. Atlas d'Anatomie Stereotaxique due Telencephale* (Masson, Cie., Paris, 1967).
15. Mintun, M. A., Fox, P. T. & Raichle, M. E. *Soc. Neurosci. Abstr.* **13**, 850 (1987).
16. Snedecor, G. W. & Corcoran, W. G. *Statistical Methods* (Iowa University Press, Iowa City, 1980).
17. Fox, P. T. *et al. Nature* **323**, 806–809 (1986).
18. Fox, P. T., Miezin, F. M., Allman, J. M., Van Essen, D. C. & Raichle, M. E. *J. Neurosci.* **7**, 913–922 (1987).
19. Damasio, A. R. & Damasio, H. *Neurology* **33**, 1573–1583 (1983).
20. Henderson, V. W. *Brain Lang.* **29**, 119–133 (1986).
21. McClelland, J. L. & Rumelhart, D. E. *Psychol. Rev.* **88**, 375–407 (1981).

22. Lauter, J., Herscovitch, P., Formby, C. & Raichle, M. E. *Hearing Res.* **20**, 199–205 (1985).
23. Mazziota, J. C., Phelps, M. E., Carson, R. E. & Kuhl, D. E. *Neurology* **32**, 921–937 (1982).
24. Roland, P., Larson, B., Lassen, N. A. & Skinhoj, E. *J. Neurophysiol.* **43**, 118–136 (1980).
25. Roeltgen, D. P., Sevush, S. & Heilman, K. M. *Neurology* **33**, 755–765 (1983).
26. Shallice, T. *Brain* **104**, 413–429 (1981).
27. Fox, P. T. & Raichle, M. E. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1140–1144 (1986).
28. Mohr, J. P. *et al. Neurology* **28**, 311–324 (1978).
29. Fox, P. T., Pardo, J. V., Petersen, S. E. & Raichle, M. E. *Soc. Neurosci. Abstr.* **13**, 1433 (1987).
30. Masdau, J. C., Schoene, W. C. & Funkenstein, H. *Neurology* **28**, 1220–1223 (1978).
31. Barris, R. W. & Schuman, H. R. *Neurology* **3**, 44–52 (1953).
32. Nielsen, J. M. & Jacobs, L. L. *Bull. L.A. Neurol. Soc.* **16**, 231–234 (1951).
33. Benton, A. L. *Neuropsychologia* **6**, 53–60 (1968).
34. Milberg, W. & Blumstein, S. E. *Brain Lang.* **14**, 371–385 (1981).
35. Milberg, W., Blumstein, S. E. & Dworetzky, B. *Brain Lang.* **31**, 138–150 (1987).
36. Humphreys, G. W. & Evett, E. J. *Behav. Brain Sci.* **8**, 689–740 (1985).
37. Shallice, T., McLeod, P. & Lewis, K. Q. *J. exp. Psychol.* **37A**, 507–532 (1985).
38. Coltheart, M., Davelaar, E., Jonasson, J. & Besner, D. in *Attention and Performance VI* (ed. Dornic, S.) (Academic, New York, 1977).

LETTERS TO NATURE

ESO400-G43, a forming galaxy?

Nils Bergvall* & Steven Jörsäter†

* Astronomiska observatoriet, Box 515, S-751 20 Uppsala, Sweden

† European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-8046 Garching bei München, FRG and Stockholms Observatorium, S-133 00 Saltsjöbaden, Sweden

Blue compact galaxies (BCGs) are characterized by their compact appearance and very blue colours^{1,2} indicative of a high star-formation rate. The metal abundances are low, suggesting that the present rate of star formation cannot have lasted for very long³ (unless the gas is being replenished). These galaxies are thus either truly young or their star formation takes place in short bursts^{3,4}. ESO400-G43 (= 2034–356) is one of the brightest and largest BCGs known. Observations in the 21-cm line with the Very Large Array (VLA) of the National Radio Astronomy Observatory reveal a massive ($5 \times 10^9 M_{\odot}$) slowly rotating H I halo. Optical observations obtained at the European Southern Observatory (ESO) show that the morphology and spectral properties are consistent with a predominantly young stellar population and an extremely low mass to light ratio ($M/L_B = 0.1$). The luminous stars seem to be concentrated in a rapidly spinning central disk. Dark matter is dominating the mass at large radii.

We have obtained charge coupled device (CCD) images of ESO400-G43 with the ESO 2.2-m telescope in the Johnson U,B,V and the *i* bands (ref. 5). An irregular appearance with numerous hot spots that show up also in the red is seen (Fig. 1). Image Dissector Scanner (IDS) spectra were obtained with the ESO 3.6-m telescope in 1980 and 1982. The spectra were corrected for IDS nonlinearity⁶. CCD spectra were obtained in 1982 and 1985 with the ESO 3.6-m telescope. The observations were corrected for extinction based on an averaged $H\alpha/H\beta$ ratio across the galaxy after correction for underlying absorption lines (the correction is $\sim 15\%$ for $H\beta$); the extinction is moderate ($A_V \approx 0.6$). The strength of the Balmer absorption lines were obtained from the spectrophotometric model discussed below.

The spectra are dominated by strong emission lines typical of H II regions heated by hot, massive stars. An electron temperature of $T_e \approx 12,000$ K was determined from the $[O III] \lambda 4,363/\lambda \lambda 4,959 + 5,007$ line ratio. We derive a gas metallicity using the ratios between the $H\beta$ and $[N II]$, $[O II]$ and $[O III]$ lines of $Z \approx 1/8 Z_{\odot}$. The blue-ultraviolet excess is large ($U-B = -0.60$; K. Olofsson, personal communication) and it is one of the intrinsically brightest objects of its type ($M_V = -20.1$). ESO400-G43 was detected by IRAS (Infrared Astronomy Satellite) at 60 μm and 100 μm .

We observed ESO400-G43 in the 21-cm H I line with the VLA in its B/C hybrid configuration in July 1985. The H I observations were made using 25 antennae with an effective channel

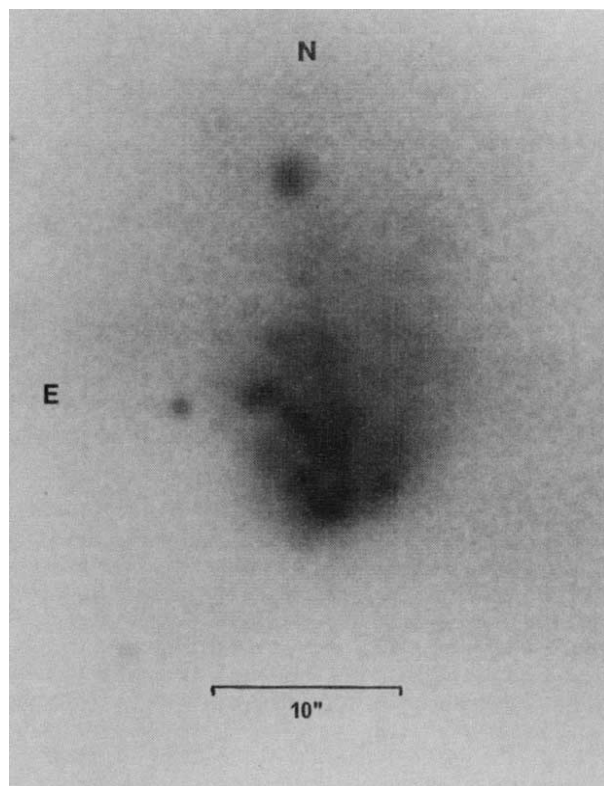


Fig. 1 A CCD image of ESO400-G43 obtained through an *i* filter ($\sim 8,200 \text{ \AA}$) with the 2.2-m telescope at ESO, La Silla, Chile. The distance is 77 Mpc assuming Hubble's constant is $75 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (1 arc s $\approx 380 \text{ pc}$).

width of 97.7 kHz corresponding to 21.0 km s^{-1} and a total number of 31 channels. The total integration time on source was 8 hours. The interferometric data were edited in the usual way and were further processed using the Astronomical Image Processing System (AIPS). The H I emission of ESO400-G43 was detected in 9 channel maps in the heliocentric velocity range $5,745\text{--}5,913 \text{ km s}^{-1}$ (optically defined).

The continuum-subtracted (using line-free channels) cleaned and smoothed (to a beamsize of $20'' \times 20''$) channel maps are shown in Fig. 2. The maps were subsequently blanked using a smoothed template in the spatial plane and afterwards interactively using the AIPS task BLANK. The resulting total H I map is shown in Fig. 3a and the derived velocity field in Fig. 3b. The total H I flux is 7.8 Jy km s^{-1} , corresponding to a total H I mass of $1.1 \times 10^{10} M_{\odot}$.

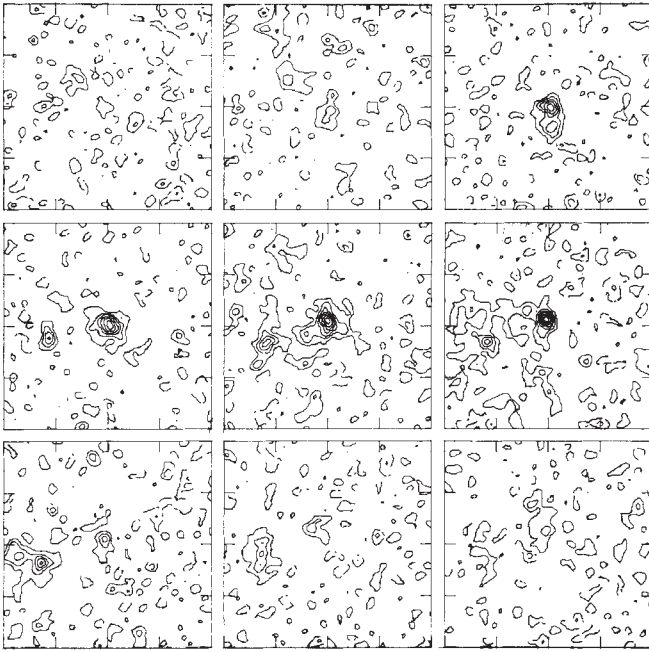


Fig. 2 The continuum-subtracted channel maps. The channel numbers increase to the right and downwards. The central image corresponds to a heliocentric (cz) velocity of $5,829 \text{ km s}^{-1}$ and the velocity interval between the channels is -21.0 km s^{-1} (velocity decreases with increasing channel number). The tickmarks are positioned at intervals of $100''$ with the central ones centred on the phase tracking point at $(\alpha = 20 \text{ h } 34 \text{ m } 31.0 \text{ s}, \delta = -35^\circ 39' 42'')$. The synthesized beam size is $20'' \times 20''$ (full width half-power, FWHP). The contours are drawn at $-2, -1$ (negative contours dashed), $1, 2, 3, 4, 5, 6$ and 7 mJy beam^{-1} .

Two major components (A and B) are evident in Fig. 3a. A is concentric with the optical image of ESO400-G43 and B is concentric with a fuzzy galaxy previously thought to be a background object. A northern feature C is also seen. The velocity is discrepant from neighbouring regions in A and the signal is low but probably real. The derived H I masses for the features are $5 \times 10^9 M_\odot$ (A), $4 \times 10^9 M_\odot$ (B) and $10^9 M_\odot$ (C). The velocity fields of the two clouds A and B suggest rotation. The position angle of the kinematic major axis (ϕ) of the A cloud is approximately 22° . This agrees well with both the position angle of the H I cloud ($\sim 15^\circ$) and that of the major axis in the *i* image ($\sim 20^\circ$). We adopt a value of 20° for the A cloud. The axial ratios ($\sim 1:1.5$) are also similar. This leads us to believe that the system is rotationally flattened. In an investigation of BCGs, Gordon and Gottesman⁷ adopt an average intrinsic axial ratio of $1:3$. Using this value we obtain an inclination of the A cloud of 50° . The observed velocity dispersion (σ_v) is $\sim 30 \text{ km s}^{-1}$ with no significant radial change.

The optical velocity field has been mapped using two CCD spectra covering the range $\lambda = 4,800\text{--}5,200 \text{ \AA}$ (full width-half maximum resolution $= 1.6 \text{ \AA}$) obtained with the slit positioned at position angle 5° and 95° through the bright spot just to the south of the centre of the galaxy in Fig. 1. The corresponding velocities are plotted in Figs 4a and 4b. Those in Fig. 4a show an unexpected behaviour. Inside 4 arc s on either side of the bright blob the velocities show a steep gradient up to about 30 km s^{-1} , after which they drop to the systemic velocity. The velocities in Fig. 4b show a similar behaviour, although the lines can only be traced out to ~ 7 arc s from the bright spot. Comparing these steep gradients to the slowly rising H I velocity field in Fig. 3b (allowing for the beam smearing), we conclude that the optical spectra outline a rapidly spinning central disk in a slower rotating H I cloud. Since $V_R/\sigma \approx 2$ (the ratio between

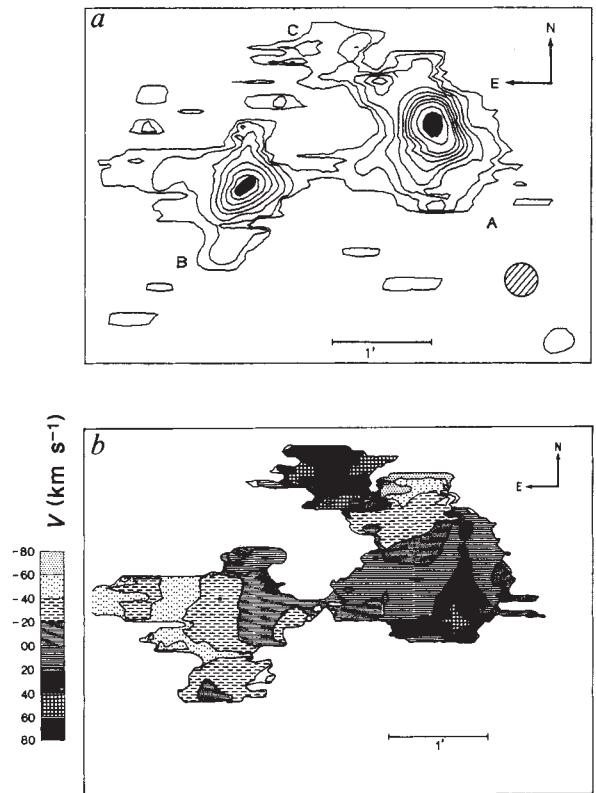


Fig. 3 A map (a) of the integrated H I distribution of ESO 400-G43. The sizes of the optical images as seen on the UK SRC B survey Schmidt plates are indicated. The contour levels are at 50, 100, 200, 250, 300, 350, 400, 450, 500 and $600 \text{ mJy beam}^{-1} \text{ km s}^{-1}$. The synthesized beam of $20'' \times 20''$ (FWHP) is indicated. The H I velocity field (b). Radial velocities relative to $5,829 \text{ km s}^{-1}$ are indicated.

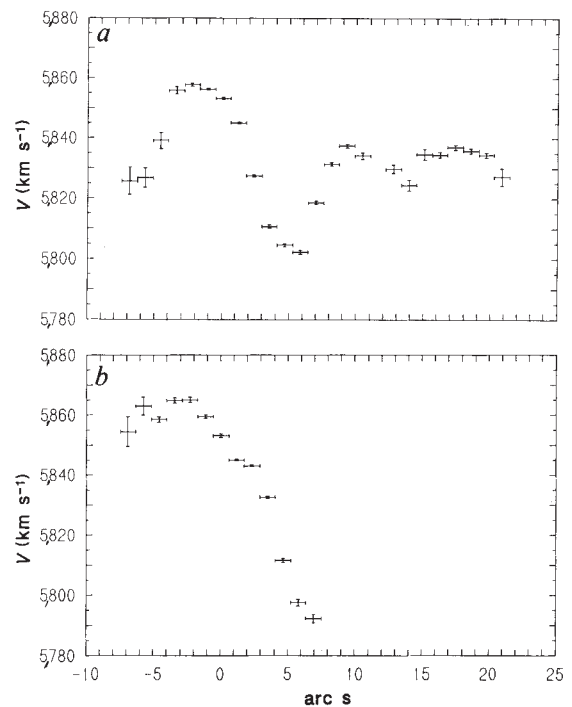


Fig. 4 The radial velocities (a) as measured from the H β and [O III] $\lambda\lambda$ 4,959 and 5,007 lines in a CCD slit spectrum at position angle 5° obtained with the ESO 3.6-m telescope at ESO, La Silla. The zero point refers to the bright spot in the southern region of the galaxy. b is the same as (a) with the slit at position angle 95° .

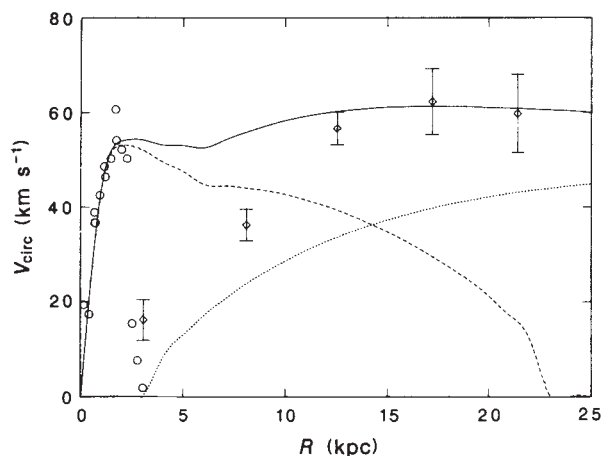


Fig. 5 Rotation velocities in ESO400-G43 projected onto the plane of the galaxy (using $\phi = 200^\circ$, $i = 50^\circ$). The circles represent the optical measurements from the spectra depicted in Fig. 4a and b. Only points situated more than 30° (in the plane of the sky) from the minor axis have been included. The rectangles with error bars show the velocities derived from the H I observations. The curves represent various fits described in the text.

rotational and random motions) we expect this disk to be rather thick; we will assume an axis ratio of 1:6.

We have modelled the stellar population and gaseous emission of ESO400-G43 through comparisons with observed spectroscopic and optical to infrared photometric properties⁸. The model is based on evolutionary tracks of stars of low metal abundances in the mass range $0.1\text{--}100 M_\odot$, constrained by the predicted equivalent width of H β in emission and the [O III]/H β ratio. The lifetime of red supergiants, a critical parameter, was estimated from statistics made on bright stars in the Small Magellanic Cloud⁹. From the evolutionary model adopting a Salpeter initial mass function¹⁰, we estimate the age of the population to be a few times 10^7 yr and the total mass in young stars and ionized gas to be $1.0 \pm 0.3 \times 10^9 M_\odot$. The predicted heavy element yield during the burst is more than adequate to account for the observed metallicity. The molecular gas mass may be derived from the IRAS fluxes ($S_{60} = 1.61$ Jy; $S_{100} = 1.51$ Jy) and the relations given by Young *et al.*¹¹; we find $M(\text{H}_2) \approx 10^8 M_\odot$. The fact that $S_{60} > S_{100}$ indicates that the dust has an unusually high temperature. Hot stars are the most likely source for the heating of the dust; this is reasonable since fluxes in the far infrared and the blue are similar.

The dynamics of ESO400-G43 may be divided into 3 regions. In Fig. 5 we have plotted the projected observed optical velocities (circles) and H I velocities (rectangles). Note that the latter are suffering from beam smearing effects in the inner parts. The steep rise and fall of the optical velocities contrast to the gradual increase of the H I velocities. The dashed rotation curve is based on the mass of stars, H II and H₂ (which dominates in the inner part) and H I (which dominates in the outer part). The fit of the central (exponential) disk with a scalelength of 0.7 kpc and an axis ratio of 1:6 yields a mass of $1.3 \times 10^9 M_\odot$. This gives $M/L_B \approx 0.1$ for the central disk. The measured scale length in the i image is 1.0 kpc, but the corrections for projection effects and extinction imply a reduction of some 30%. The fit to the optical velocity data is quite good up to the maximum. After the maximum the optical velocities show a sharp non-keplerian drop. This is presumably due to the fact that these parts of the disk are not yet relaxed. The optical spectra seem to rule out an increase in the velocity dispersion that fully compensates for the velocity drop.

The outer part of the dashed curve is dominated by the model exponential H I spheroid with an axis ratio of 1:3. The best fit to the observed (spatial) distribution is a scale length in the plane of 4 kpc. A pressure term of the form $P = 1/3 \sigma_v^2 \rho$ is

included, and it is in fact the dominant term over most of the range in radius. The measured velocities differ from the calculated in the outer parts. An isothermal dark halo with a core radius of 8 kpc and $\sigma = 33 \text{ km s}^{-1}$ has been fitted (dotted line) to account for the discrepancy. The solid line shows the resulting rotation curve. The rotation curve can be traced out to more than 15 optical scalelengths of the disk.

The galaxy ESO400-G43 seems to be a unique object from several aspects. We think that ESO400-G43 is presently forming its first massive generation of stars. No older stellar component is required to model the emission or kinematic properties of ESO400-G43. It is almost impossible to rule out completely the existence of a small old population, but we have shown that it is possible to explain the observed properties of the stellar component without invoking old stars. We can only speculate what triggered the star formation in this galaxy. A likely explanation is interaction with the relatively massive companion. The fact that ESO400-G43 galaxy seems to contain dark matter is interesting, especially since the 'visible' mass is dominated by H I and we can thus infer the existence of dark matter without any references to M/L ratios.

We wish to thank the VLA staff, and one of us (N.B.) wishes to thank ESO for its hospitality in October 1985 during which a part of this work was completed. VLA is operated by the National Radio Astronomy Observatory. This work was partly supported by the Swedish National Science Research Council.

Received 20 August; accepted 30 November 1987.

1. Sargent, W. L. *Astrophys. J.* **160**, 405–427 (1970).
2. Sargent, W. L. & Searle, L. *Astrophys. J.* **162**, L155–L160 (1970).
3. Searle, L. & Sargent, W. L. *Astrophys. J.* **173**, 25–33 (1972).
4. Gerola, H. & Seiden, P. E. *Astrophys. J.* **223**, 129–139 (1978).
5. Wade, R. A., Hoessel, J. G., Elias, J. H. & Huchra, J. P. *Publ. astr. Soc. Pacif.* **91**, 35–40 (1979).
6. Rosa, M. *The Messenger* **39**, 15–17 (1985).
7. Gordon, D. & Gottesman, S. T. *Astr. J.* **86**, 161–177 (1981).
8. Bergvall, N. & Olofsson, K. *Astr. Astrophys. Suppl.* **64**, 469–476 (1986).
9. Humphreys, R. M. & McElroy, D. B. *Astrophys. J.* **284**, 565–577 (1984).
10. Salpeter, E. E. *Astrophys. J.* **121**, 161–167 (1955).
11. Young, J. S., Schloerb, F. P., Kenney, J. D. & Lord, S. D. *Astrophys. J.* **304**, 443–458 (1986).

Planetographic clustering of low-altitude impulsive electric signals in the night ionosphere of Venus

C. T. Russell, M. von Dornum & F. L. Scarf

Institute of Geophysics and Planetary Physics,
University of California, Los Angeles, California 90024, USA

The possible occurrence of lightning in the Venus atmosphere has long been a subject of speculation^{1,2}, so the initial reports of impulsive electromagnetic signals deep in the Venus atmosphere³ and in the night ionosphere⁴ generated much interest. The results of optical experiments, however, were mixed. The infrared spectrometer on Venera 9 observed one sequence of lightning-like activity⁵. The star-sensor of the PVO orbiter⁶ saw none and the Vega balloon photometers⁷ obtained negative but ambiguous results. Evidence for the lightning generation of the electromagnetic signals includes their impulsive nature, their occurrence patterns^{8,9}, their altitude dependences^{8,10} and their electromagnetic polarization¹¹. Although only signals below several hundred hertz can directly propagate to the PVO spacecraft, higher frequency signals could scatter in the inhomogeneities of the night ionosphere¹². As these higher frequencies cannot propagate far into the ionosphere and are seldom observed at altitudes of above 400 km (ref. 13) they mark the location of the generating region. In this paper we map the locations of the observations made at 730 Hz during the first three night-time observing seasons of Pioneer Venus, when the spacecraft reached altitudes low enough to detect

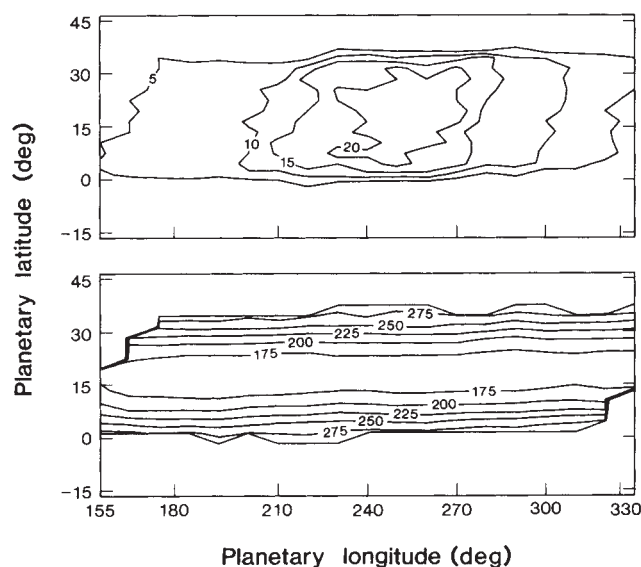


Fig. 1 Contour maps of number of observations per 3° latitude by 10° longitude bin (upper panel) and average altitude in a bin (lower panel) as a function of Venus longitude and latitude.

these leakage signals. These maps show that there are specific 'active regions' of enhanced signal occurrence and that at least some of these active regions appear to remain fixed in planetographic coordinates from year to year.

The Pioneer Venus orbiter carries a simple plasma-wave instrument with an electric field antenna whose output is filtered and recorded in four narrow channels at 0.10, 0.73, 5.4 and 30.0 kHz (ref. 4). With the usual magnetic field strength in the night ionosphere (10–30 nT) and the typical plasma density at the peak of the ionosphere ($\sim 1.000 \text{ cm}^{-3}$), electromagnetic waves can propagate directly to the spacecraft from electrical discharges only at 100 Hz. Such signals have been extensively studied by Scarf and co-workers^{4,9–11}. Whistler-mode waves are guided only loosely by the magnetic field and hence extrapolations of the waves to their source locations may be inaccurate. We study the nightside ionosphere and not the dayside because in the dayside ionosphere the magnetic field is usually weaker than on the nightside and when it is strong the field is horizontal and radial, thus inhibiting propagation of any electromagnetic waves out of the atmosphere. Moreover, the noise level of the Pioneer Venus electric field detector is much higher when the spacecraft is in direct sunlight.

Electromagnetic waves at 0.73, 5.4 and 30 kHz should not propagate freely in the night ionosphere: any such signals attempting to propagate into the ionosphere from below should be efficiently reflected back into the ionosphere below. But this reflection process is not instantaneous and the wave energy can propagate some distance into the ionosphere. The patchy nature of the electron density in the night ionosphere aids this process¹². As these higher frequency signals cannot propagate far from their source, they are good tracers for any source region that may be present. On the negative side, as these waves can be detected only at the very lowest altitudes only a very limited portion of the Venus surface can be probed. This limits the latitude range and the number of observing seasons that can be studied. The altitude of periapsis of Pioneer Venus was controlled to maintain it at low altitudes for only the first three Venus years. Fortunately, because of the closeness of Venus's diurnal and orbital periods, the region on the surface of Venus covered each season had some overlap, allowing us to check for repeatability of the occurrences from year to year. Because the same basic pattern is seen at all three frequencies, 0.73, 5.4 and 30.0 kHz, in this note we discuss only one of them, the 0.73-kHz data.

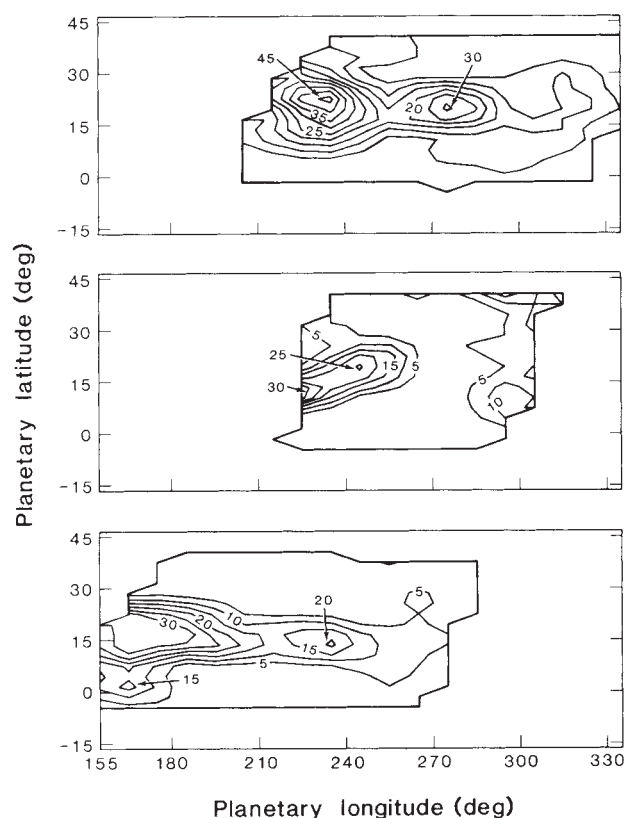


Fig. 2 Contour maps of the rate of occurrence of 730-Hz burst activity as a function of planetary longitude and latitude for the three night-time low-altitude observing seasons.

We have chosen to characterize every 30 s of Pioneer Venus observations of plasma waves in the night ionosphere as having either 'bursty' activity or not having it. Bursty activity is considered to be any signal whose change in strength in 30 s is comparable with or greater than the mean and whose amplitude rose above a threshold of $1.5 \times 10^{-5} \text{ V mHz}^{1/2}$. We have sorted these observations into bins according to the latitude and longitude of the subsatellite point. Because the spacecraft moves in latitude $\sim 3^\circ$ in 30 s we have chosen the latitudinal width of our bins to be 3° . The longitudinal width of each bin was chosen to be 10° so that a reasonable number of observations would be accumulated in each bin as the satellite moved overhead. Typically each observing season contributed about seven observations to a bin, each observed on a successive day. Thus the points within each bin can be considered independent of each other. The points in adjacent bins in the north-south direction are not necessarily independent samples because the spacecraft moves across the planet from north to south.

The upper panel of Fig. 1 shows the number of observations in each bin over the region mapped. The observations on the far right-hand side were obtained in season 1 and on the far left-hand side in season 3. In the middle of the map, data are available for all three seasons. Only data at solar zenith angles $> 90^\circ$ and within 0.95 Venus radii of the Sun-Venus line were used in this study.

The lower panel of Fig. 1 shows the average altitude of the observations in each bin. The PV spacecraft reached periapsis at a latitude of $\sim 15^\circ \text{ N}$ during this period. Both above and below this latitude the average altitude of the spacecraft climbed rapidly. The altitude of the spacecraft at periapsis was maintained at $\sim 150 \text{ km}$ by the expenditure of gas.

The three panels of Fig. 2 show the occurrence rates of 730-Hz bursts in each of the three seasons separately. In season 1 there was an active region to the west of Beta at $\sim 270^\circ \text{ E}$ longitude

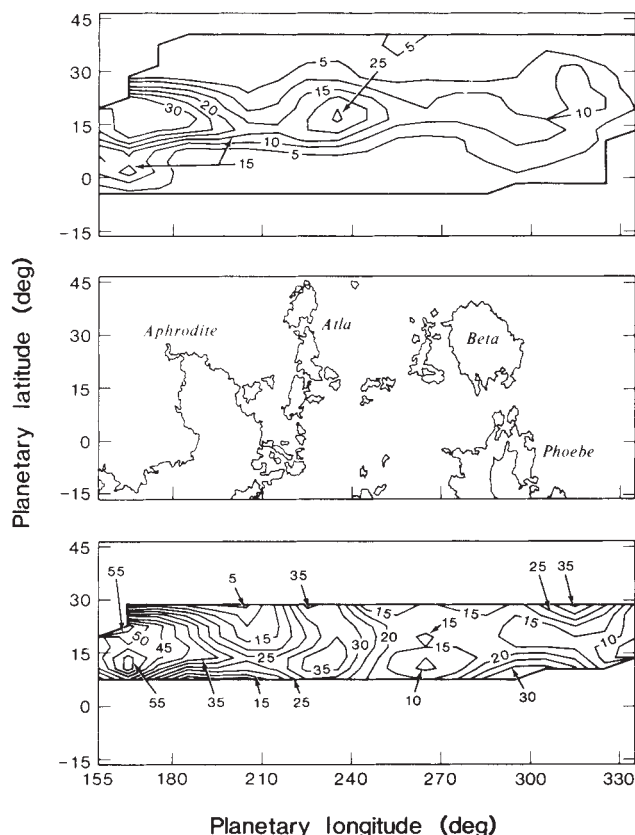


Fig. 3 Contour maps of occurrence of 730-Hz activity averaged over the three seasons (top) and corrected for the altitude dependence of the occurrence of the signals (bottom). The middle panel shows Venus highland regions as outlined by the 6,052.5 km altitude contour.

and one over Atla at 230° E. In season 2 there was an enhancement over Phoebe at 295° E and one over Atla at 230° E. In season 3 there were no data over Beta and Phoebe but the active region over Atla at 230° E was present and a new active region was seen at 180° E north and west of Aphrodite. In short, the active region over Atla was present in all three seasons, although variable in strength, whereas the active region to the west of Beta is transitory. These sites are close to regions identified as being associated with 100-Hz bursts in earlier studies.

The top panel of Fig. 3 shows the occurrence rate for the three seasons combined. The occurrence rate of bursts is quite high in the region of peak occurrence, >25% above Atla and >30% above the northwest Aphrodite active region. If the source of these bursts is lightning then lightning must be occurring nearly continually in the Venus atmosphere at some of these sites. The top panel of Fig. 3 also shows that the peak activity seems to be concentrated at ~15° N latitude, the periapsis latitude of Pioneer Venus. The season for this concentration is the rapid decrease in amplitude and hence occurrence of these signals with altitude.

Figure 4 illustrates the rapid decrease in occurrence rate with altitude. In this plot we have calculated the occurrence rate over all longitudes and latitudes as a function of altitude in both 10-km and 50-km bins. Examining first the 50-km bins, we see that the occurrence rate levels off at about 2% at high levels. We consider this to be our noise level in this analysis. If we correct our 50-km occurrences by subtracting the high-altitude occurrence rate we obtain the three open circles in Fig. 4. The straight line shows our fit to this corrected radial decrease. Its slope corresponds to a scale length of 40 km. This length is ~10% of the wavelength of 730-Hz electromagnetic waves in free space. The small points show the occurrence rates in 10-km

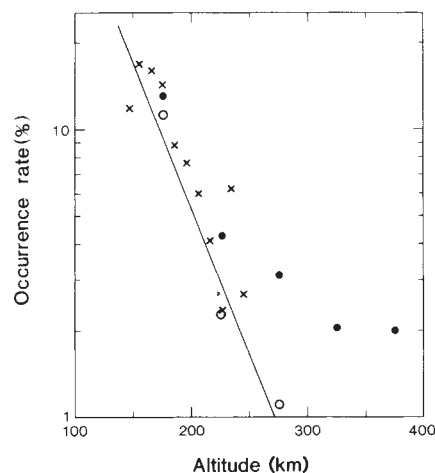


Fig. 4 The altitude dependence of the occurrence rate of 730-Hz bursts. Occurrences have been calculated in 10-km (crosses) and in 50-km (filled circles) bins as appropriate and have been corrected for the background rate of occurrence of such bursts as seen at high altitudes (open circles).

bins. They follow the same trend but have greater variability because they are more affected by the occurrence of individual strong active regions. In the available data set it is not possible to explore the altitude dependence of the occurrence rate over a single active region. At lowest altitudes the rate of activity appears to drop. This effect appears to be due to a change in the gain of the electric field antenna when the spacecraft enters the denser layers of the lower ionosphere¹³.

In the lower panel of Fig. 3 we use this average altitude variation to normalize all occurrence rates to 150 km. We believe we can separate latitudinal from altitudinal variations because we sample a wide range of topographic features at each altitude and because closest approach is at 15° N and not at the Equator. Had closest approach been over the Equator, latitudinal variations that were symmetrical about the Equator would be inextricably mixed with the altitude variation.

We do not expect Venus to be north-south symmetrical about 15° N. We make this correction only for bins whose average altitude is <225 km to avoid errors associated with large correction factors. The correction process has two quite noticeable effects. First, the occurrence rates are no longer maximum at 15° N. Second, the sources over Atla and northwest of Aphrodite have shifted more toward the Equator. The maximum occurrence rate over these sites has increased to 35% and 55%, respectively. In the eastern region on the map there is a strong active (maximum occurrence rate 30%) to the east of Beta. There is also a strong activity (maximum occurrence 30%) centred on Phoebe.

With the limited overlap available for three seasons it is not possible to prove unambiguously that each of the features in the occurrence rate maps was time-stationary or was just transiently present, but the strong activity over Atla was present for all three seasons. The strong activity over Phoebe was present for both seasons in which there were observations. We feel that most of the maxima in occurrence rate shown in Fig. 3 are probably permanent features associated with weather patterns, or storms. The fact that some active regions are close to but not right over highlands implies that volcanoes or tall mountains are probably not the immediate cause of the lightning. This agrees with the conclusions of Borucki¹⁴, who argues on the basis of a lack of aerosols at low altitudes and the high dielectric strength of CO₂.

The variation of occurrence rate from season to season suggests that other factors such as local time also affect the occurrence rate. In this regard we note that observations over a fixed planetary longitude occur at a different local time each season.

We thank Peter Ford for providing the Venus topography data used in constructing the map in Fig. 3. This work was supported by NASA.

Received 16 September; accepted 6 November 1987.

1. Meinel, A. B. & Hoxie, D. T. *Comm. Lunar Planet. Lab., Univ. Arizona* **1**, 35-38 (1962).
2. Davidson, G. T. & Anderson, A. D. *Science* **156**, 1729-1730 (1967).
3. Ksanfomaliti, L. V. *Kosm. Issled.* **17**, 747-762 (1979).
4. Taylor, W. W. L., Scarf, F. L., Russell, C. T. & Brace, L. H. *Science* **205**, 112-114 (1979).
5. Krasnopol'skii, V. A. *Kosm. Issled.* **18**, 429-434 (1980).
6. Borucki, W. J., Dyer, J. W., Thomas, G. Z., Jordan, J. C. & Comstock, D. A. *Geophys. Res. Lett.* **8**, 233-236 (1981).
7. Sagdeev, R. Z. *et al. Science* **231**, 1411-1414 (1986).
8. Ksanfomaliti, L. V., Scarf, F. L. & Taylor, W. W. L. in *Venus* (eds Hunten, D. M., Colin, L., Donahue, T. M. & Moroz, V. I.) 565-603 (Univ. Arizona, Tucson, 1983).
9. Scarf, F. L. & Russell, C. T. *Geophys. Res. Lett.* **10**, 1192-1195 (1983).
10. Scarf, F. L., Jordan, K. F. & Russell, C. T. *J. geophys. Res.* **92**, 12407-12411 (1987).
11. Scarf, F. L. & Russell, C. T. *Science* (in the press).
12. Singh, R. N. & Russell, C. T. *Geophys. Res. Lett.* **13**, 1071-1074 (1986).
13. Russell, C. T., von Dornum, M. & Scarf, F. L. *J. geophys. Res.* (in the press).
14. Borucki, W. J. *Icarus* **52**, 354-364.

Observations of the sodium layer at high latitudes in summer

U. von Zahn, G. Hansen & H. Kurzawa

Institute of Physics, University of Bonn, 5300 Bonn 1, Nussallee 12, FRG

At altitudes between 80 and 100 km the upper atmosphere contains a layer of free sodium atoms presumably maintained by the almost continuous inflow and vaporization of cosmic dust particles. This sodium layer has been studied by ground-based photometers and lidar (laser radar), rocket-borne instrumentation, and satellite-borne photometers. Sodium measurements have been carried out at low, middle and high geographic latitudes. Lidar observations show that the peak of the layer commonly occurs close to the altitude of 90 km, with maximum densities of $\sim 5,000$ atoms cm^{-3} . In the 1960s photometric measurements established that the sodium layer shows seasonal variations which become more pronounced towards high latitudes (for a review of these early measurements, see ref. 1). Later lidar measurements yielded the following ratios of midsummer over midwinter column densities: 4.5 to 1 at 40° N (ref. 2), 3 to 1 at 44° N (ref. 3), and 6 to 1 at 51° N (ref. 4). Polar latitudes, however, are constantly sunlit in summer and the measurement of Na density profiles under these particular conditions have not yet been possible by ground-based methods. The state of our theoretical understanding of the Na layer has recently been reviewed by Kirchhoff⁵. Here we report on results of the first lidar measurements of the Na layer performed at polar latitudes in summer. Our sodium lidar instrument located at the Andøya Rocket Range, Norway (69° N, 16° E) gave measurements over ten days during July 1986 and July 1987. We find the sodium layer confined entirely to the lower thermosphere. The mean values for the layer parameters are as follows: the sodium column density is $6 \times 10^8 \text{ cm}^{-2}$; the number density at the layer peak is $1.2 \times 10^3 \text{ cm}^{-3}$; the altitude of the layer peak is 93 km, and the halfwidth of the layer is 6 km. Comparing these values with previous lidar measurements made at the same site in winter we obtain the following ratio of winter to summer densities: 10 to 1 for the sodium column density and 5 to 1 for the sodium number density at the layer peak.

The lidar instrument used for our measurements employs an excimer laser, a narrowband dye laser working at 589 nm, and a 1-m-diameter receiving telescope. The instrument is described in greater detail in ref. 6. To enable this instrument to make daylight measurements we decreased the divergence of the laser beam and the field of view of the receiving telescope to 0.25 mrad (full-width at half-maximum) and 0.36 mrad, respectively. In

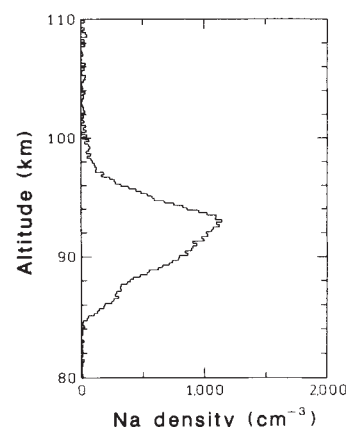


Fig. 1 Sodium profile averaged over ~ 17 h of observations between 15 July 1987, 07:40 UT and 16 July 1987, 22:04 UT. Individual altitude channels are 200-m wide. Noise peaks outside the sodium layer are due to the background in the raw data caused by sunlight.

addition to that, Fabry-Perot interferometers have been placed in front of the photon-counting multiplier so that an effective receiver bandwidth of either 15 or 30 pm is provided for. The instrument is located near Andenes, Norway (69.3° N, 16.0° E). For ten days atmospheric measurements were carried out for a total of 56 h. These observations cover all local times. The integration period for individual profiles was chosen between 2 and 10 min. The altitude resolution (channel width) was set to 200 m.

Figure 1 shows a mean sodium profile, averaged over data obtained between 15 July 1987, 07:41 UT and 16 July 1987, 22:04 UT (local standard time is UT + 1 h). The following features of this profile are of major interest and are typical of most of the rest of our July profiles: the Na density at the peak of the layer is only $\sim 1,000 \text{ cm}^{-3}$, the altitude of the layer peak is near 93 km, and the bottom of the layer is near 85 km. Thus, it appears that at polar latitudes free sodium atoms are found in summer only in the thermosphere.

The distribution of our observations both over the months of July and over local time are sufficiently even to make the calculation of monthly mean values of the most important layer parameters meaningful. From the July measurements we have so far calculated Na profiles by assuming the appropriate CIRA 1972 (*COSPAR International Reference Atmosphere 1972*, Akademie-Verlag, Berlin, 1972) temperature profile (needed to evaluate the Na resonance cross-section) and by setting the optical thick-

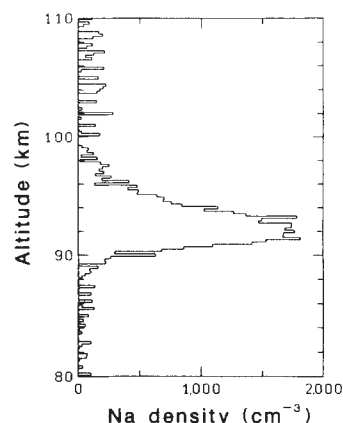


Fig. 2 Sodium profile averaged over 2 h of observations between 01 and 03 UT on 26 July 1987. Additional comments as in Fig. 1.

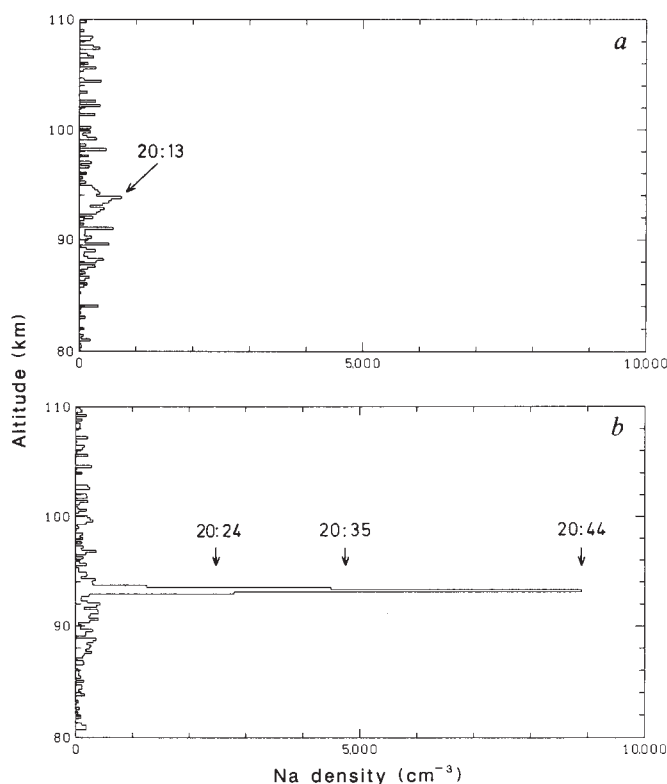


Fig. 3 Sodium profiles averaged over 5 min of observations on 24 July 1987. The profile shown in (a) was obtained about 20:13 UT, the profile of (b) about 20:44 UT. The halfwidth of the layer at the latter time is close to 400 m (or just two of our altitude channels). The peak sodium densities at two intermediate times are also given in (b).

ness of the Na layer to zero. For July and 69° N latitude we obtain the following monthly mean values: Na column density = $(6 \pm 2) \times 10^8 \text{ cm}^{-2}$; Na density at the layer peak = $1,200 \pm 300 \text{ cm}^{-3}$; altitude of the layer peak = $93 \pm 1 \text{ km}$; halfwidth of the layer = $6 \pm 1 \text{ km}$.

The error bars quoted reflect the current and as yet preliminary state of our data analysis (they are not meant to indicate genuine geophysical variability). We expect these error bars to decrease as soon as more elaborate data processing has been completed. Nevertheless, even with their current error bars these values provide interesting new information because the two density parameters deviate significantly from values found at other latitudes.

Figure 2 demonstrates a feature of Na profiles frequently found at our observation site: a sharp cutoff of the layer at its bottom side. The example shown was taken from data obtained on 26 July 1987, averaging over the time period 01 to 03 UT. Even for this average over 2 h the cutoff altitude is still well defined. In the course of a day, however, cutoff altitudes might vary by a few km, thus smearing out this signature in long-term averages.

As discovered by von Zahn *et al.*⁷ and described in greater detail by von Zahn and Hansen⁸ the normal sodium layer above Andøya is frequently enhanced by the rapid formation of additional and very narrow Na layers, the halfwidth of which is $\sim 1 \text{ km}$. Figure 3 gives an impressive example taken from our recent observation campaign and measured under sunlit conditions on 24 July 1987. Figure 3a shows a weakly developed Na layer in the 92–95 km altitude region. This profile was obtained after 5 min of integration time, centred on 20:13 UT. In the following half hour a very strong and extremely narrow Na layer developed at 93.5 km altitude. Figure 3b shows the profile at 20:44 UT at which time the Na density had reached its

maximum of $9,000 \text{ cm}^{-3}$ (peak densities at two intermediate times are also indicated in Fig. 3b). Even after 30 min of growth the layer still exhibits the extremely small halfwidth of 400 m. The stability of the very small scale heights, observed at the bottomside and topside of the layer for more than half an hour, seem to indicate that very low turbulence persisted in this atmospheric region at the time of this measurement. von Zahn *et al.*⁷ suggested that these sudden Na layers are produced *in situ* by the impact of energetic auroral particles upon a sodium-containing atmospheric trace compound. It is evident that the occurrence of these events makes the definition of 'average' properties of the Na layer somewhat ambiguous. Here we do not intend to deal with this subject matter in more depth. We state that in calculating the mean parameters quoted above, these sudden layers have been considered part of the Na layer (and hence have not been removed in any way from the profiles before averaging them).

The outstanding result is the observation of very low Na densities and, to a lesser extent, of an unusual layer profile. We first compare the values we measured in summer with those measured in winter at the same site by von Zahn and Tilgner⁹. The latter determined the following mean values for the period from December to February: column density $5.8 \times 10^9 \text{ cm}^{-2}$; density at the layer peak $6.1 \times 10^3 \text{ cm}^{-3}$; altitude of the layer peak 89.2 km, and halfwidth of the layer 9.8 km. Thus, in comparison with the winter values the summer peak density is reduced by a factor of 5 and the layer halfwidth by a factor of 2. The combined effect of these two features is to reduce the column density by a factor of 10. In addition, the altitude of the layer peak is $\sim 4 \text{ km}$ higher in summer than in winter.

On 1 August 1971 Witt *et al.*¹⁰ measured two twilight sodium profiles by means of rocket-borne photometers at 68° N latitude. They found the Na layer to be rather narrow (halfwidth $< 3 \text{ km}$) and with a sharp bottomside near 87 km. Both these features are confirmed by our new lidar results. Yet, Witt *et al.*¹⁰ obtained a Na density at the layer peak of $10,000 \text{ cm}^{-3}$, which clearly differs from our mean value, but was derived neglecting self-absorption and multiple scattering effects (G. Witt, personal communication).

Summer lidar measurements now provide for the following dependence of Na column densities on latitude: close to $3 \times 10^9 \text{ cm}^{-2}$ at 23° S (ref. 11), 40° N (ref. 2), and 44° N (ref. 3); $2.5 \times 10^9 \text{ cm}^{-2}$ at 51° N (ref. 4) and $0.6 \times 10^9 \text{ cm}^{-2}$ at 69° N (this work). A very recent result of Gardner *et al.*¹² indicates a column density of $0.6 \times 10^9 \text{ cm}^{-2}$ also at 78° N.

The feebleness of the Na layer observed at high latitudes in summer is probably related to enhanced loss of sodium atoms through photoionization^{13,14} and to the extremely low temperatures in the mesopause region under these conditions. These low temperatures will, for example, lead to a decrease in the Na densities as a result of chemical reactions having rate constants with inverse temperature dependence¹⁵. Other contributing factors, not yet studied systematically are an enhanced adsorption probability of Na atoms on smoke, dust and cloud particles, a change in the trace gas composition brought about by the mean upward bulk motion of the atmosphere under the geophysical conditions studied here, or perhaps new ion chemistry caused by very low temperature which may lead to an increased loss of free sodium to sodium-bearing ions. Polar mesospheric cloud particles and noctilucent cloud particles probably absorb Na atoms. They are known, however, to exist only near the bottom of the summer Na layer. Furthermore, even an infinitely strong sink for sodium atoms at the bottom of the sodium layer will not prevent *per se* the formation of a normal layer (ref. 13). We argue that smoke particles such as those proposed by Hunten *et al.*¹⁶ could provide a powerful sink for Na atoms not only close to the bottomside of the Na layer but throughout the sodium layer, in particular at low temperatures. Clearly, additional observations, especially of the ambient temperature profile, are much needed to advance our

understanding of the Na layer at high and polar latitudes in summer.

We thank the staff of the Andøya Rocket Range for excellent support of our activities, P. von der Gathen for assistance in processing the raw data, and D. M. Hunten for helpful comments on the manuscript. This research is supported by Deutsche Forschungsgemeinschaft, Bonn, Germany.

Received 24 August; accepted 3 November 1987.

- Hunten, D. M. *Space Sci. Rev.* **6**, 493–573 (1967).
- Gardner, C. S., Voelz, D. G., Sechricht, Jr, C. F. & Segal, A. C. *J. geophys. Res.* **91**, 13659–13673 (1986).
- Megie, G., Bos, F., Blamont, J. E. & Chanin, M. L. *Planet Space Sci.* **26**, 27–35 (1978).
- Gibson, A. J. & Sanford, M. C. W. *J. atmos. terr. Phys.* **33**, 1675–1684 (1971).
- Kirchhoff, V. W. J. H. *Can. J. Phys.* **64**, 1664–1672 (1986).
- von Zahn, U. & Neuber, R. *Beiträge Phys. Atmos.* **60**, 294–304 (1987).
- von Zahn, U., von der Gathen, P. & Hansen, G. *Geophys. Res. Lett.* **14**, 76–79 (1987).
- von Zahn, U. & Hansen, T. *J. atmos. terr. Phys.* (in the press).
- von Zahn, U. & Tignier, C. *Eur. Space Ag. Symp. Proc. SP-270*, 107–112 (1987).
- Witt, G., Stegman, J. & Wood, H. in 'Methods of Measurements and Results of Lower Ionosphere Structure' (ed. K. Rawer) 431–445 (Akademie-Verlag, Berlin, 1974).
- Simonich, D. M., Clemesha, B. R. & Kirchhoff, V. W. J. H. *J. geophys. Res.* **84**, 1543–1550 (1979).
- Gardner, C. S., Senft, D. C., Kwon, K. H. & Sechricht, C. F. Jr. *Nature* (submitted).
- Gadsden, M. *Ann. Geophys.* **26**, 141–150 (1970).
- Hunten, D. M. *Geophys. Res. Lett.* **8**, 369–372 (1981).
- Swider, W. *Geophys. Res. Lett.* **12**, 589–591 (1985).
- Hunten, D. M., Turco, R. P. & Toon, O. B. *J. atmos. Sci.* **37**, 1342–1357 (1980).

Structure of $(\text{CuO})_2$ double layers in superconducting $\text{YBa}_2\text{Cu}_3\text{O}_7$

H. W. Zandbergen, R. Gronsky, K. Wang & G. Thomas

National Center for Electron Microscopy, Materials and Chemical Sciences Division, Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720, USA

Since the report that multiphase mixture with the nominal composition $\text{La}_{2-x}\text{Ba}_x\text{CuO}_{4-y}$ exhibit possible superconductivity¹, research on this class of quaternary oxides has increased enormously. The superconducting compound in the original mixture was rapidly identified as $\text{La}_{1.85}\text{Ba}_{0.15}\text{CuO}_4$ ², and subsequent research concentrated on the most suitable cation substitutions in this compound. A new oxide with a considerably higher critical temperature, T_c , was very soon found in the phase system $\text{Y}_2\text{O}_3\text{--BaO--CuO}$ ³, later identified by several groups as $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ ⁴, and now commonly known as the 1–2–3 compound. Intensive activity in structural characterization followed including reports of various planar defects. Here we report an analysis, using high-resolution electron microscopy of planar defects in superconducting $\text{YBa}_2\text{Cu}_3\text{O}_7$, revealing the identity of a CuO double layer intercalation. Within this double layer, the Cu atoms remain in planar fourfold coordination with O, identical to their arrangement in the single CuO layers of the $\text{YBa}_2\text{Cu}_3\text{O}_7$ matrix.

The structure of the 'bulk' material has been precisely determined by neutron powder diffraction^{5–8} in agreement with single crystal X-ray diffraction⁹, and is shown in Fig. 1A. High-resolution electron microscopy (HREM) has been performed by several groups^{9–14}, all confirming the 'bulk' structure, but also reporting the existence of several types of local structural modifications, mostly planar in character.

Ourmazd *et al.*¹³ report the existence of two extra layers (one Y layer and one CuO layer) inserted at the position of the CuO layer ($z=0$). Viegars *et al.*¹⁵ report defects at the grain boundary having a lattice parameter of 1.38 nm instead of the 1.16 nm characteristic of the undeformed structure. Antiphase boundaries with a $c/3$ shift were characterized by Zandbergen *et al.*¹⁶ and Domenges *et al.*¹⁷. Another type of planar defect was found¹⁶ to exist at the mirror plane locations ($z=0$ or $z=1/2$) in $\text{ErBa}_2\text{Cu}_3\text{O}_7$. The defect, an inclusion of an extra layer leading to a c -axis dimension of ~ 1.36 nm, was shown to be a double Er layer with a stacking order similar to the double LaO layer

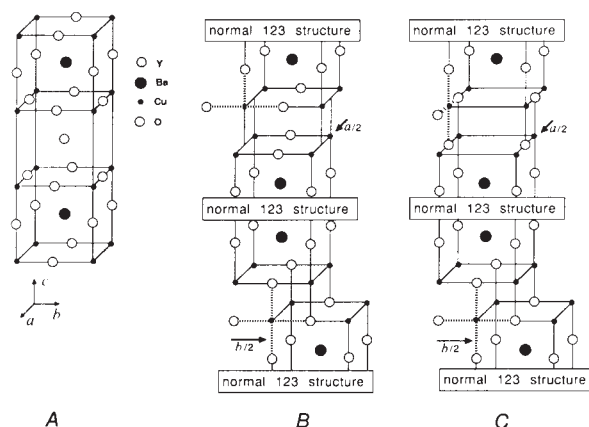


Fig. 1 The structure of perfect $\text{YBa}_2\text{Cu}_3\text{O}_7$ (A) and the defect layers for two different configurations (B) and (C), used in image calculations. Note that the shift along the a -axis is accompanied by a rearrangement of the oxygen atoms in (C).

in $\text{La}_2\text{Sr}_x\text{CuO}_4$ ². Furthermore $[100]$ 90° rotation twins, across which the b -axis aligns with a c -axis, have also been studied¹⁸.

Recently, the decomposition of the 1–2–3 phase was observed to nucleate at free surfaces and proceed by the insertion of an extra CuO layer at the location of the original CuO single layer¹⁹. This insertion leads to an enlargement of the c -axis by 15%, due to the presence of $(\text{CuO})_2$ double layers. The decomposition reaction occurs even at room temperature and may be enhanced by moisture. It is emphasized that in freshly cleaved specimens there is no evidence of planar defects resulting from inserted CuO layers. Consequently many of the early HREM reports on inclusions and stacking defects along the c -axis are questionable because at that time, it was not known that the material decomposed upon contact with air.

The purpose of this research is to clarify the nature of the $(\text{CuO})_2$ double-layer defects in superconducting $\text{YBa}_2\text{Cu}_3\text{O}_7$ by detailed local structure determination employing high-resolution electron microscopy and computer-assisted image analysis. High-resolution electron microscopy was carried out in a JEOL JEM 200CX electron microscope operating at 200 kV and equipped with a top entry $\pm 10^\circ$ double-tilt goniometer, and the Berkeley atomic resolution microscope²⁰ operating at 1,000 kV and equipped with a $\pm 40^\circ$ biaxial, double-tilt-lift goniometer.

Specimens for high-resolution electron microscopy were prepared in a variety of ways. Powdered $\text{YBa}_2\text{Cu}_3\text{O}_7$ was cooled quickly from room temperature to 77 K in liquid N_2 , and a few droplets of a suspension of the powder were put on to conventional porous carbon films over Cu grids. In a parallel study, small wedge-shaped fragments (~ 0.5 mm long) were selected after cleaving pieces of sintered pellets of $\text{YBa}_2\text{Cu}_3\text{O}_7$. These fragments were mounted in Cu or Au folding grids. Some material was also subjected to ion milling; however, these samples were found to contain a prominent amorphous layer which complicates image interpretation and in addition showed a marked increase in their rate of defect formation. Because this study is sensitive to such defect formation, the ion milling procedure was abandoned.

Image calculations were performed using the CEMPAS simulation software developed at the National Center for Electron Microscopy²¹. Image processing was carried out using SEMPER software²² on images digitized via an Eikonix camera.

In Fig. 2, two different images of defects are observed, as indicated by the different contrast features within the planar-faulted regions (arrowed). Computer simulation was used to determine whether these images came from two different defects or whether both images came from the same defect in different orientations. The procedure is relatively straightforward, as the

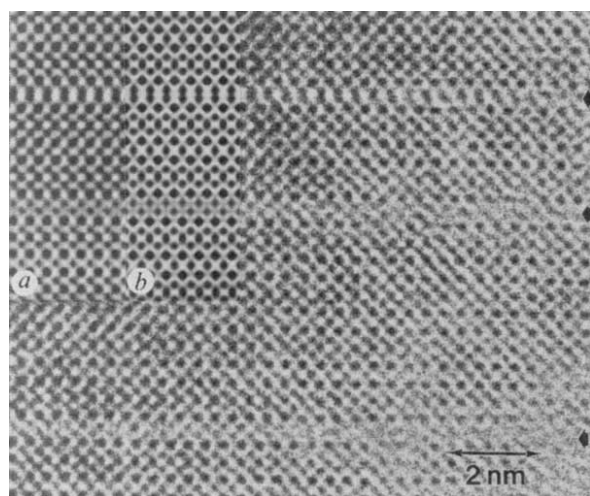


Fig. 2 HREM image of $\text{YBa}_2\text{Cu}_3\text{O}_7$ in [010] orientation showing defects (indicated by arrows) formed at room temperature. Two different defect images can be seen. The upper defect shows mirror symmetry and the other two, glide symmetry. Note that the upper defect gradually assumes glide character near the edge, suggesting that the length of the a -axis is different on both sides of the defect. An averaged experimental image (a) and a calculated image (b) are inset.

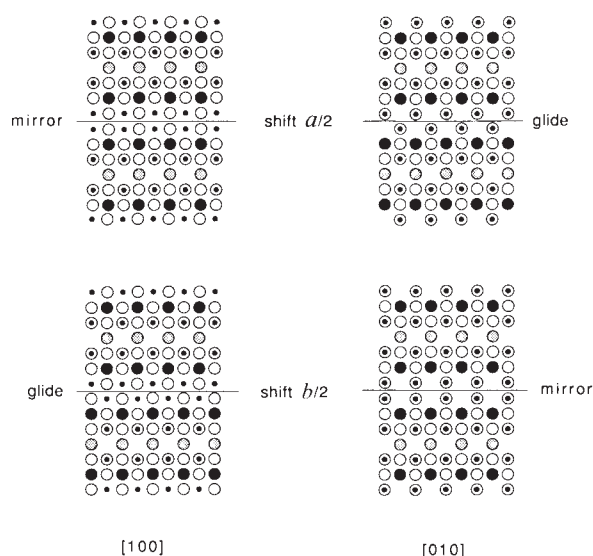


Fig. 3 Structure models showing the resulting projected images with a shift of $a/2$ and $b/2$. A shift of $a/2$ leads to a defect layer with mirror symmetry when projected along [100] and glide symmetry when projected along [010]. For a $b/2$ shift, the symmetries are reversed. The key for atomic representation is the same as in Fig. 1.

surrounding image of the perfect $\text{YBa}_2\text{Cu}_3\text{O}_7$ matrix serves as an internal calibration on specimen thickness and the defocus. Images of this 'normal' structure were first simulated using published lattice parameters of $\text{YBa}_2\text{Cu}_3\text{O}_7$ (ref. 23). Matching between experimental and computed images was sought by varying the objective lens defocus between 0 and -140 nm and the specimen thickness between 2 and 20 nm in the calculations. Images of defect structures were then calculated using a computational 'unit cell' with a c -axis seven times larger than that of the actual 1-2-3 structure. In this model, two defect layers were introduced as shown in Fig. 1B: one with a shift between the CuO planes of $a/2$ (henceforth called the $a/2$ defect) and one with a shift of $b/2$ (the $b/2$ defect). It is noted that, depending upon the viewing or projection direction, these defects show either mirror or glide symmetries (see Fig. 3). The distance between the two defects was taken to be two standard unit cells apart, a distance observed in one of the HREM images. Calculations were carried out for four different double CuO layers in which the occupancy of the Cu sites was either $1/2$ or 1 and oxygen was situated at the normally occupied or normally vacant positions of the lattice, as shown in Fig. 4.

To quantify the intensity variation in the images, line scans were made across the defects in the averaged experimental images and the calculated images, shown for comparison in Fig. 4. From these scans it is evident that the intensity of all Cu atoms is identical except for those Cu atoms within the defect layer.

Using the image of the $\text{YBa}_2\text{Cu}_3\text{O}_7$ matrix as a reference, it was found that the best fit between the experimental images and the calculated ones occurred when the calculational model was orientated in an [010] direction. The distinguishing feature in these images is the fact that the intensities of all Cu atom positions are approximately the same; this is to be expected as they all appear in identical mixed columns (Cu and O) along the [010] projection direction. In the orthogonal [100] projection, the Cu atoms at the $z = 0$ plane should image with reduced intensity because they are not interspersed with oxygen atoms along the projection direction (this is clearly seen in the calculations). Applying the same comparison technique to the defect images, it is seen that the defect showing mirror symmetry fits the model well in an [010] orientation while the defect showing glide symmetry fits the model better in the wrong, that is, the

orthogonal [100], orientation. This error necessitates a change in the proposed model.

An obvious correction is suggested by the image-matching results obtained in this study; namely, the observed contrast always appears to be consistent with double $\text{Cu-O-Cu-O} \dots$ chains along the projected image direction. If the atomic arrangement of the $a/2$ defect is made the same as in the proposed model but with locally reversed a - and b -axes (a configuration that can be achieved by exchanging oxygen atoms and vacancies in the $z = 0$ plane), then the resulting structure will have double $\text{Cu-O-Cu-O} \dots$ chains that are aligned in the direction of the shift (along the a -axis). If the displacement that is necessary to accommodate the inclusion of the extra CuO layer occurs along the b -axis, then the structure of the $\text{YBa}_2\text{Cu}_3\text{O}_7$ matrix is retained, again with double $\text{Cu-O-Cu-O} \dots$ chains along the direction of the shift (the b -axis). We used this model in detailed calculations for image matching.

Image simulations were carried out with a slice thickness of 0.382 and 0.389 nm for the [100] and [010] directions, respectively. Under these conditions all the atoms in the unit cell are projected onto one plane and any information regarding the position of the atoms within the unit cell along the electron beam direction is ignored. Although this approximates the experimental conditions obtained in high-resolution phase-contrast imaging, it prohibits any determination from a single image whether the observed glide defect pattern of Cu atoms is due to a shift of $a/2$ alone, or whether the shift contains an orthogonal component (for example, $a/2 + ab$ where α may assume any fractional value). Fortunately such information can be obtained by comparing high-resolution electron micrographs obtained in different specimen orientations. For the defects imaged here, an obvious direction to compare with the [100] image is the [110] image; together these two projections make it possible to decide clearly whether the contrast observed at the location of the $(\text{CuO})_2$ double layer is due to a displacement along one axis or two.

Using the high-angle-tilting capability of the Berkeley atomic resolution microscope, it was possible to image the same defective regions of a crystal in both [100] and [110] orientations at high resolution. In these experiments, it was observed that in $\langle 110 \rangle$ projections, the images revealed a shift of $0.5d_{(110)}$ at the defect sites. Furthermore mirror and glide defect patterns are

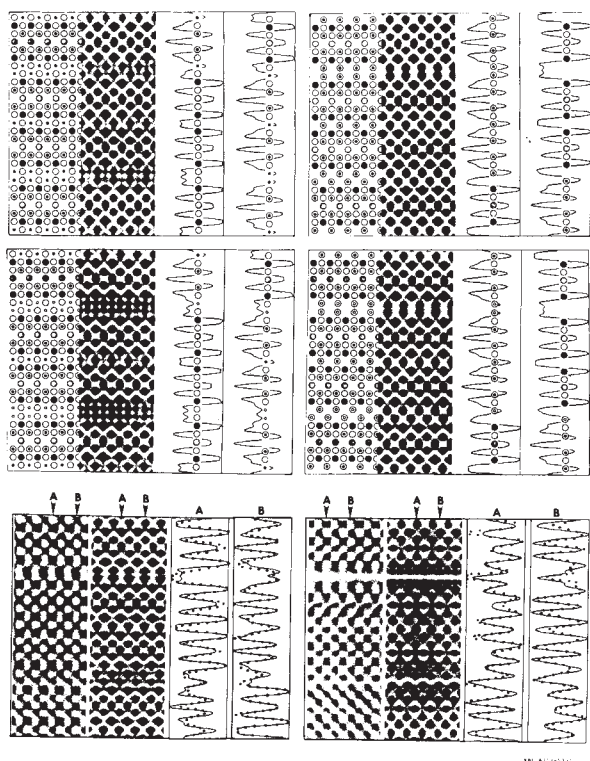


Fig. 4 Two different interpretations of the structure viewed along [100] and [010]. A projection of the structure model, a calculated image, and line scans are shown in *a-d*. Full small circles represent fully occupied Cu positions and open small circles represent half-occupied Cu positions. The locations at which the scans were made are indicated by the schematic row of atoms on which they are superimposed, where the thickness of a scanned line was 0.075 nm. The averaged experimental and calculated images at two different defocus values are shown in *e* and *f*. Full lines represent calculated data, dotted lines the observed intensity variations in the images. The parameters used in the calculations are: specimen thickness 2 nm; objective aperture 6 nm^{-1} ; acceleration voltage 1,000 kV; spherical aberration coefficient 2.8 mm; beam divergence angle 0.6 mrad.

observed for both the [100] and [010] orientations. These observations suggest that introduction of the extra CuO layer coincides with a shift of $a/2$ or $b/2$ leading to glide and mirror defect features in the images as shown in Fig. 2.

From the line scans in Fig. 4, it is clear that the contrast of the Cu and O atoms in the mirror defect is nearly identical to the contrast at the Y-atom position. This implies that we cannot determine directly from the images whether the black dots in the mirror defect originate from a mixed column of Cu and O atoms or from a column of Y atoms. However, a number of chemical and structural considerations narrow the possibilities.

Direct substitution of Cu and O atoms by Y atoms would necessitate Y having twofold coordination with oxygen. This follows from the high-resolution results; the observed intensity variations are only consistent with isolated Y ions in the double layer, not Y and O, because the mixed atomic column image would lead to different intensity variations. However, twofold coordination is not an acceptable configuration for Y, because Y requires at least sixfold coordination with oxygen. Furthermore, insertion of an Y atom at the Cu position would increase the c -axis lattice parameter, due to the larger size of the Y^{3+} ion. Direct measurements of the high-resolution images show that the actual increase of the c -axis due to the insertion of the defect layer is $15.0 \pm 0.5\%$, whereas values of 17% or more are to be expected if Y is included in the defect plane. Moreover, the observed 15% increase in the c -axis indicates that the Cu-O distance in the defect plane along the c -axis is ~ 0.185 nm, which

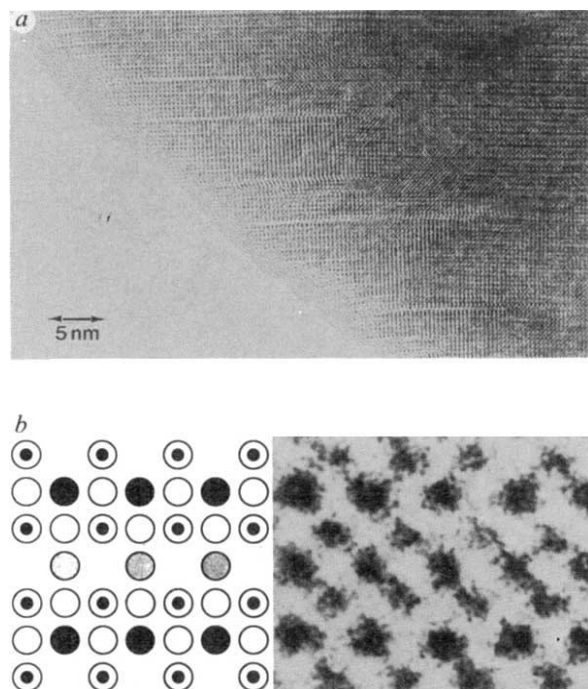


Fig. 5 *a*, HREM image of the edge of a crystal in [100] or [010] orientation showing that the insertion of the CuO planes leads to a bending of the lattice. Often this insertion event is observed to occur in pairs. Due to a difference in the amount of bending along the electron beam direction between the surface layers and the interior layers, the symmetry of the image is locally reduced. *b*, HREM image of $\text{YBa}_2\text{Cu}_3\text{O}_7$ in [010] orientation. The columns of Ba, Y and Cu are seen as separate black dots; atoms in the model are labelled as in Fig. 1.

is exactly the Cu-O distance along the c -axis for the $\text{YBa}_2\text{Cu}_3\text{O}_7$ matrix observed by neutron powder diffraction²⁰.

The best fit between experimental and calculated images was obtained for the model in which Cu-O-Cu-O... chains occur in the direction of the displacement required to accommodate the extra CuO layer. This means that a shift along the b -axis will leave the oxygen atoms at their original position, whereas a shift along the a -axis will lead to a rearrangement of the oxygen atoms in both CuO planes (Fig. 2). In the latter case the originally vacant sites will be occupied by oxygen atoms. Only in this way can the planar fourfold coordination of Cu be maintained. This planar fourfold coordination is often observed in Cu compounds (for instance CuO), and is known to be energetically favourable. That it is also a stable coordination in this specific structure can be deduced from the fact that it exists in the 'bulk' $\text{YBa}_2\text{Cu}_3\text{O}_7$ matrix and furthermore, for $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ in the approximate range $0.4 < \delta < 0$, a superstructure occurs, marked by a doubling of the a -axis and an alternating sequence of Cu-O-Cu-O... and Cu-vacancy-Cu-vacancy... chains having fourfold and twofold coordination of Cu atoms respectively¹⁰.

This study also reveals a helpful feature of phase-contrast imaging applied to the oxide superconductors. Because of their very slight differences in image contrast, it is usually very difficult to distinguish between a $\text{YBa}_2\text{Cu}_3\text{O}_7$ crystals in [100] orientation or in [010] orientation, particularly because the rotational symmetry of both directions is the same. Two methods can be used however. First, the c/a ratio can be determined. Unfortunately, for $\text{YBa}_2\text{Cu}_3\text{O}_7$ this is not very reliable because the lattice can expand in a direction perpendicular to the c -axis (see Fig. 1) in thin regions. Furthermore, one must be aware that the measured c/a ratio will also depend on imperfections in the lenses (of both the electron microscope and the photographic projector)

and the position of the image on the photographic negative (deviations generally increase towards the edge of the negative).

Second, comparisons between calculated and observed images can be used. Once more it is unfortunate that the difference between images calculated for the [100] and [010] orientations is so small, as the structural differences between these directions arise solely from the oxygen/vacancy ordering in the CuO plane. However, as is shown above, it is possible to determine the crystal orientation by distinguishing the contrast due to the presence or absence of oxygen ions at the $z = 0$ plane, provided the crystal is thin enough and the point resolution is such that isolated Cu atoms can be imaged separately (see Fig. 5a). Although this is a stringent set of requirements, they do identify conditions under which the [100] or [010] directions of the orthorhombic phase can be distinguished.

The most serious problem in the interpretation of all defects is the presence of local lattice bending in the neighbourhood of their core structures as is shown in Fig. 5b. When the crystal is very thin (≤ 4 nm) and the defect is extended through the entire sample thickness, the best conditions for imaging are obtained. However, if the defect is restricted to the surface or if it jogs over neighbouring CuO planes, which will occur often in thicker crystals, the lattice will bend to accommodate the structural changes of the defect. The image obtained from such configurations will show local differences in contrast, very similar to bend contours in thin crystals. Due to such local changes in the lattice orientation, direct interpretation of the image is impossible. Image simulations could, in principle, be used to determine the structure, but calculations are typically very time-consuming and do not guarantee any single choice of actual structure.

Although there are no data yet available on the effect of (CuO)₂ double layers on superconducting performance, it would be interesting to note in particular how the critical current varies with the density of such defects, as they cause an increase in the number of superconduction electrons (that is, those within the CuO planes). A possible problem may be the bifurcation of the conduction planes at the junction between a double layer and the normal single-layer structure, and any determination of the atomic structure of the bifurcation region will also require imaging in more than one orientation. If each CuO layer is replaced by a (CuO)₂ double layer a superstructure is obtained. Shifts along the b -axis in all (CuO)₂ double layers lead to a structure with space group $Ammm$ ²⁴. When all the shifts are along the a -axis the same structure is obtained because of the rearrangement of the oxygen atoms in the (CuO)₂ double layer. Such a regular superstructure has indeed been observed in thin films^{24,25}. In contrast with bulk material no strong bending of the lattices is observed, making it the perfect material to study the influence of the double layer on the superconductivity.

This work is supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Division of Materials Sciences, US Department of Energy. The technical assistance of the staff of the National Center for Electron Microscopy is gratefully acknowledged.

Received 11 January; accepted 27 January 1988.

1. Bednorz, J. G. & Müller, K. A. *Z. Phys.* **B64**, 189–193 (1986).
2. Jorgensen, J. D. *et al. Phys. Rev. Lett.* **58**, 1024–1027 (1987).
3. Wu, M. K. *et al. Phys. Rev. Lett.* **58**, 908–910 (1987).
4. Grant, P. M. *et al. Phys. Rev. B* **35**, 7242–7244 (1987).
5. Tarascon, J. M., Greene, L. H., McKinnon, W. R. & Huli, G. W. *Phys. Rev. B* **35**, 7115–7118 (1987).
6. Hazen, R. M. *et al. Phys. Rev. B* **35**, 7238–7241 (1987).
7. Siegrist, T., Sunshine, S., Murphy, D. W., Cava, R. J. & Zahurak, S. M. *Phys. Rev. B* **35**, 7137–7139 (1987).
8. Cox, D. E., Moodenbaugh, A. R., Hurst, J. J. & Jones, R. H. *J. Phys. Chem. Solids* (in the press).
9. Garbaskas, M. F., Arendt, R. H. & Kasper, J. S. *Inorg. Chem.* **26**, 3191–3194 (1987).
10. Zandbergen, H. W., van Tendeloo, G., Okabe, T. & Amelinckx, S. *Phys. Status Solidi (a)* **103**, 45–72 (1987).
11. Takeda, S. & Hikami, S. *Jap. J. appl. Phys.* **26**, L848–L851 (1987).
12. Beyers, R. *et al. Appl. Phys. Lett.* **50**, 1918–1921 (1987).
13. Ourmazd, A. *et al. Nature* **327**, 308–310 (1987).
14. Marks, L. D., Tsuruta, D., Theil, J. & Poeppelmeier, K. in *Proc. 45th A. Meet. Electron Microscopy Soc. Am.* (ed. G. W. Bailey) 56–57 (San Francisco Press, San Francisco, 1987).

15. Vieggers, M. P. A. *et al. J. Mater. Res.* **2**, 743–749 (1987).
16. Zandbergen, H. W., Holland, G. F., Tejedor, P., Gronsky, R. & Stacy, A. M. *Adv. Ceramic Mat.* **2**, 688–701 (1987).
17. Domenges, B., Hervieu, M., Michel, C. & Raveau, B. *Europhys. Lett.* **4**, 211–214 (1987).
18. Zandbergen, H. W. *et al. Proc. Mater. Res. Soc., Fall Meeting, Boston* (1987).
19. Zandbergen, H. W., Gronsky, R. & Thomas, G. *Phys. Status Solidi (a)* (in the press).
20. Gronsky, R. & Thomas, G. in *Proc. 41st A. Meet. Electron. Microscopy Soc. Am.* (ed. Bailey, G. W.) 310–311 (Claitor's, Baton Rouge, 1983).
21. Kilaas, R. in *Proc. Microbeam Analysis Society* (in the press).
22. Saxton, W. O. *Inst. Phys. Conf. Ser.* **44** (Inst. Phys., London, 1978).
23. Torardi, C. C. *et al. Am. chem. Soc. Symp. Ser.* **357**, 152–163 (1987).
24. Marshall, A. F. *et al.* preprint.
25. Zandbergen, H. W. & Thomas, G. *Phys. Status Solidi (a)* (submitted).

Structure of K⁺(cryptand[2.2.2]) electride and evidence for trapped electron pairs

R. H. Huang, M. K. Faber*, K. J. Moeggenborg, D. L. Ward & J. L. Dye†

Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, USA

Electrides are crystalline salts in which stoichiometric amounts of trapped or itinerant electrons serve as the anions^{1–8}. The cations are alkali metal cations complexed by cyclic or bicyclic polyethers of the crown ether⁹ or cryptand¹⁰ classes. Optical spectra, powder conductivities and magnetic susceptibilities show that, in most electrides, individual electron localization occurs, presumably centred at the anionic sites, with trap depths of 0.5–1.0 eV. An exception is K⁺(cryptand[2.2.2]) · e[−], which has a plasma-like optical absorption spectrum², high microwave conductivity, a low activation energy for direct current powder conductivity (~0.02 eV), and a weak, temperature-dependent electronic paramagnetic susceptibility. The crystal structure of this electride shows the presence of large (4 × 6 × 12 Å) vacancies of complex shape, interconnected in two directions by zigzag channels, but blocked in the third direction. The structure and properties suggest the presence of weakly bound electron pairs.

The first indication that electrides might exist was the observation of “blue, strongly paramagnetic solids”¹¹ at the time of preparation of the first crystalline salt of an alkali metal anion, Na⁺(cryptand[2.2.2]) · Na[−]. A similar observation was reported by Harris and Lagowski¹². The optical spectra of thin electride films^{1,2} show that most electrides have a distinct absorption peak in the near-infrared, characteristic of electron localization and similar to that of solvated and trapped electrons¹³. However, films made by rapid evaporation of ammonia from solutions

* Present address: Corning Glass Works, Corning, New York 14831, USA.

† To whom correspondence should be addressed.

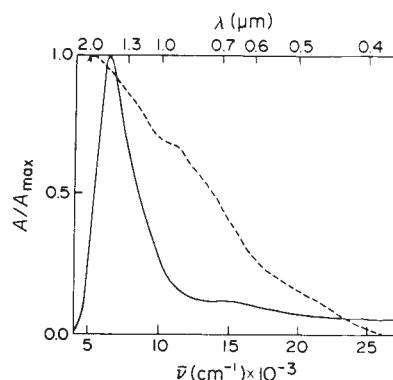


Fig. 1 Optical absorption spectra of solvent free electride films: solid line, Cs⁺(18C6)₂ · e[−], dashed line, K⁺(C222) · e[−].

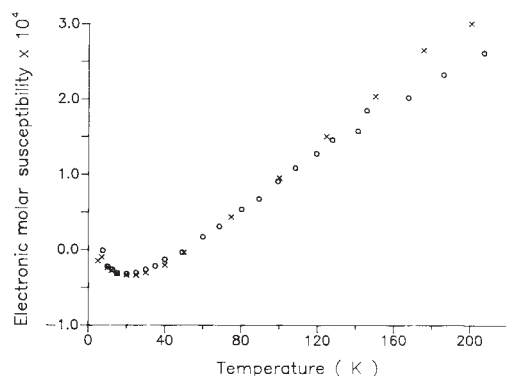


Fig. 2 Molar electronic magnetic susceptibility of $K^+(C222) \cdot e^-$. Obtained from the overall susceptibility by subtracting the susceptibility of the decomposed sample (diamagnetic contribution) and the low-temperature Curie-law paramagnetism. The symbols X and O represent separate runs on different preparations.

that contained 1:1 mixtures of potassium and cryptand[2.2.2] (abbreviation C222) or 2:1 mixtures of lithium and cryptand[2.1.1] showed strikingly different spectra², with continually rising absorbance into the near-infrared, very similar to the spectra of metallic solutions of the alkali metals in ammonia¹³. This suggested either electron delocalization or the presence of shallow traps. The spectra of thin films of $Cs^+(18C6)_2 \cdot e^-$ and $K^+(C222) \cdot e^-$ shown in Fig. 1 illustrate these differences.

Polycrystalline samples of the title compound were prepared by dissolving stoichiometric amounts of potassium metal and cryptand[2.2.2], in dimethyl ether at -50°C , adding a less polar solvent (trimethylamine or diethyl ether) and cooling slowly to -78°C . The rigorously anaerobic preparation vessel, the vacuum-line techniques used and the apparatus for measuring d.c. conductivities have been previously described⁵. Single crystals were grown from a dimethyl ether-diethyl ether mixture that contained seed crystals of $K^+(C222) \cdot e^-$ by cooling from -50°C to -67°C over a period of two days. The shiny, black, plate crystals were transferred to a Nicolet P3F diffractometer in a cold nitrogen stream and kept at -70°C during data collection.

Earlier studies of the microwave power absorption of powders produced by rapid solvent evaporation, and the spectra of thin films, suggested that $K^+C222 \cdot e^-$ is metallic². The high microwave conductivity was confirmed by measuring the power absorbed by polycrystalline samples contained in 2.0-mm inner diameter fused silica tubes in a TE₁₀₃ X-band microwave cavity. The response with most electrifieds was essentially the same as that of an empty tube, while that of the title compound was similar to the response with finely divided metal powders. Direct current powder conductivity measurements over the temperature range -170 to -120°C showed, however, a positive temperature coefficient of conductivity, with an activation energy of ~ 0.02 eV, suggesting that this compound is not metallic, but rather contains weakly bound electrons. Alternatively, the positive temperature coefficient could arise from grain boundary effects.

The electronic contribution to the magnetic susceptibility of the title compound is shown in Fig. 2. A low-temperature Curie paramagnetism, resulting from 0.5–1.0% localized unpaired electrons (preparation-dependent) has been subtracted. The residual diamagnetism at low temperatures shows that most electron spins are paired as the temperature approaches 0 K. The increase in susceptibility with temperature results from either thermally induced electron-pair dissociation or population of a triplet state. Sample decomposition limits the measurements to temperatures $< -70^\circ\text{C}$ (at which temperature the susceptibility is $\sim 14\%$ of the value expected for completely uncoupled electrons) so that only an approximate value of the pairing

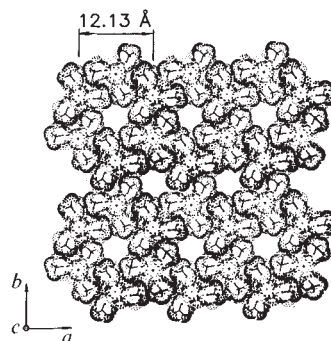


Fig. 3 View down the z -axis of the complexed cation packing in the electrified $K^+(C222) \cdot e^-$. This view was constructed from the X-ray data by using an Evans and Sutherland PS-300 display system and appropriate van der Waals radii of the atoms. The hexagonal holes represent channels in the z direction of diameter ~ 4 Å.

energy (0.05–0.1 eV) can be obtained. It may be noted that solvated electron spin-pairing also occurs in solution^{13,14}.

The optical, electrical and magnetic behaviour of $K^+(C222) \cdot e^-$ is thus very different from that of most electrifieds and indicates weak electron-pair trapping. The locale of such trapping is suggested by the crystal structure. X-ray diffraction studies at $-71 \pm 2^\circ\text{C}$ of a single crystal of dimensions $0.2 \times 0.6 \times 0.8$ mm yielded the structural data given in Table 1 (monoclinic space group $C2/c$). The structure of the complexed cation is normal, with K^+ –O distances ranging from 2.795 to 2.858 Å, K^+ –N distances of 2.958 and 2.981 Å, and the usual interatomic distances in the complexant¹⁵. It is the packing of the complexed cations and the resulting 'empty' channels and vacancies that make the structure interesting.

Computer-generated displays of the structure that use van der Waals radii to represent atomic sizes illustrate the molecular packing. The view down the z -axis (nearly along c), shown in Fig. 3, demonstrates that the strand of one cryptand molecule nests into the gap between strands on its neighbours. The effect is to form nearly hexagonal patterns with channels at the centre of each group of six cryptated cations. The z -coordinates of the K^+ ions alternate around the ring as in the chair form of cyclohexane with a vertical separation of 1.7 Å. Views perpendicular to the y -axis through the centres of the z and x channels, respectively, are shown in Fig. 4. The two sets of channels intersect to form large, dumb-bell-shaped cavities of approximate dimensions $4 \times 6 \times 12$ Å, with $\sim 4 \times 4$ Å cross-section at the centre. The minimum diameter of the z channels is ~ 4 Å, defined by cryptand hydrogens with internuclear distances of ~ 6.4 Å across the channels. The roughly rectangular ($\sim 4 \times 5$ Å) channels along x are constricted midway between their intersections with the z channels by two hydrogens on opposite sides whose

Table 1 Crystal data for $K^+(C222) \cdot e^-$

Space group	Monoclinic $C2/c$
Cell parameters	$a = 12.129(8)$, $b = 20.692(13)$, $c = 21.519(16)$ Å $\beta = 95.23(6)^\circ$, $V = 5,378(6)$ Å ³ , $Z = 8$
No. of reflections collected	6,692
No. of unique reflections	6,394
No. of reflections used in refinement	3,614 with $F_o^2 > 3.0\sigma(F_o^2)$
Maximum 2θ	45°
Unweighted agreement factor	0.041
Weighted agreement factor	0.040
High peak in final diff. map.	$0.10(1)$ e Å ⁻³
Low peak in final diff. map	$-0.11(1)$ e Å ⁻³

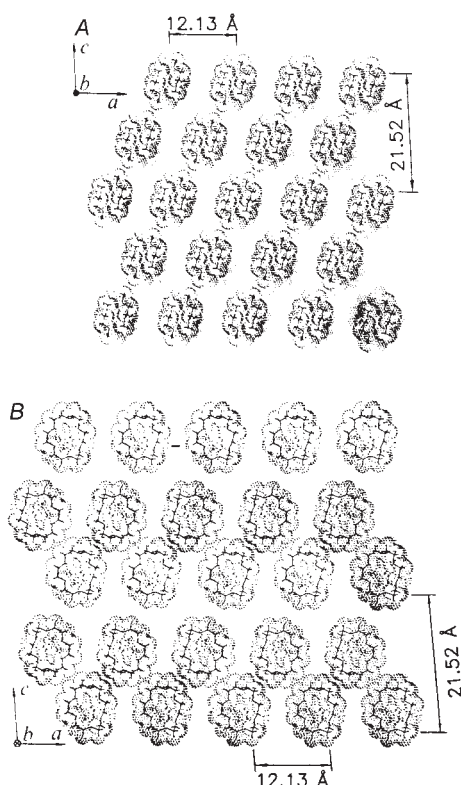


Fig. 4 Cutaway views down the y -axis of the channels in $K^+(C222) \cdot e^-$: **A**, through the centres of the z axis channels, which form the openings shown in Fig. 3; **B**, through the centre of one set of x -axis channels. View **B** is displaced by ~ 1.5 Å along y from that shown in **A**; displacement in the opposite direction would show the other set of x -axis channels. The apparent blockage of the x -axis channel in this view is not real, but results from alternating small displacements of the channel along the y direction. This view illustrates the formation of the dumbbell-shaped electron-pair trapping centres at the intersections of the x and z channels.

centres are ~ 4 Å apart. There are no appreciable channels along the y direction, giving the structure a distinctly two-dimensional character. The crystal cleavage planes are parallel to the directions of the channels.

Possible electron-pair trapping sites are the large cavities at the intersections of the x and z channels. In the rubidide, $Rb^+(C222) \cdot Rb^-$, which has comparable cation packing, pairs of Rb^- ions occupy similar sites. In contrast to the electride $Cs^+(18\text{-crown-6})_2 \cdot e^-$, which is isostructural⁶ with the corresponding sodide, $Cs^+(18\text{-crown-6})_2 \cdot Na^-$, the structure of $K^+(C222) \cdot e^-$ is very different from that of $K^+(C222) \cdot Na^-$ (orthorhombic, $Fdd2$, $a = 15.769$, $b = 25.245$, $c = 13.818$ Å). In the latter compound the Na^- ions are trapped individually at sites that are isolated from each other (unpublished results).

We conclude that the structure and properties of the electride, $K^+(C222) \cdot e^-$ are consistent with the presence of electron pairs in a singlet ground state, weakly trapped in large vacancies that are interconnected by channels to form a two-dimensional network. The effective volume occupied by the electron pairs is probably larger than the vacancy size given above, because we expect considerable overlap with the cryptand hydrogens that line the channels. Studies of single-crystal conductivities, ESR spectra and single-crystal reflectivities are in progress and should tell us more about the nature of electron-electron interactions in this compound.

We acknowledge support from the NSF and the Michigan State University Center for Fundamental Materials Research. We thank Mark Kuchenmeister for assistance with display of the crystal structures.

Received 7 September; accepted 7 December 1987.

1. Dye, J. L., Yemen, M. R., DaGue, M. G. & Lehn, J.-M. *J. chem. Phys.* **68**, 1665–1670 (1978).
2. DaGue, M. G., Landers, J. S., Lewis, H. L. & Dye, J. L. *Chem. Phys. Lett.* **66**, 169–172 (1979).
3. Ellaboudy, A., Dye, J. L. & Smith, P. B. *J. Am. chem. Soc.* **105**, 6490–6491 (1983).
4. Dye, J. L. & Ellaboudy, A. *Chem. Brit.* **20**, 210–215 (1984).
5. Dye, J. L. *Prog. inorg. Chem.* **32**, 327–441 (1984).
6. Dawes, S. B., Ward, D. L., Huang, R. H. & Dye, J. L. *J. Am. chem. Soc.* **108**, 3534–3535 (1986).
7. Dye, J. L. & DeBacker, M. G. A. *Rev. Phys. Chem.* **38**, 271–301 (1987).
8. Dye, J. L. *Scient. Am.* **257**, 66–75 (1987).
9. Pedersen, C. J. *J. Am. chem. Soc.* **89**, 7017–7036 (1967).
10. Dietrich, B., Lehn, J.-M. & Sauvage, J. P. *Tetrahedron Lett.* 2885–2888 (1969).
11. Tehan, F. J., Barnett, B. L. & Dye, J. L. *J. Am. chem. Soc.* **96**, 7203–7208 (1974).
12. Harris, R. L. & Lagowski, J. J. *J. phys. Chem.* **82**, 729–734 (1978).
13. Thompson, J. C. *Electrons in Liquid Ammonia* (Oxford University Press, 1976).
14. Edwards, P. P. *Adv. inorg. Chem. Radiochem.* **25**, 135–185 (1982).
15. Moras, P. D., Metz, B. & Weiss, R. *Acta crystallogr.* **B29**, 383–388 (1973).

The siting, energetics and mobility of saturated hydrocarbons inside zeolitic cages: methane in zeolite Y

Subramanian Yashonath, John M. Thomas*,
Andreas K. Nowak & Anthony K. Cheetham

Davy Faraday Research Laboratory, The Royal Institution,
21 Albemarle Street, London W1X 4BS, UK

To understand how reactant hydrocarbons are catalytically converted into more useful products inside the cages of zeolite solids, it is first necessary to evaluate the strength of binding and ease of migration of the reactant within the cages, the inner walls of which house the active sites. As it is not easy to arrive at the properties by direct experiment, especially at elevated temperatures, it is profitable to consider computer simulation, which we show to be a valuable alternative approach for the case of methane inside a model, faujasitic catalyst ($Na_{48}Al_{48}Si_{144}O_{384}$, abbreviated Na-Y). The information that can be gleaned from the potential-energy distribution functions evaluated by Monte Carlo methods^{1,2} in the range 10–298 K augurs well for the computational study of bulkier reactants inside other zeolitic catalysts.

Our appreciation of the factors that control the energetics of adsorption and diffusion of hydrocarbons in zeolites is rather poor; this is surprising in view of the obvious technical significance of these systems. There is a wealth of thermodynamic data from which heats of adsorption can be calculated, but the range of reported experimental values for a particular system is often disturbingly wide. The task of retrieving reliable data is complicated because for just one type of zeolitic framework, based say on faujasite, the mineral progenitor of zeolites X and Y (Fig. 1), numerous structural variants are possible^{3–5}: for each Si/Al ratio, which itself can take up one of ten or so discrete values between 1 and 3, there are several ways of distributing the Al in the tetrahedral framework and, consequently, of accommodating the counterbalancing exchangeable cation in

* To whom correspondence should be addressed.

Table 1 Parameters used for the calculation of short- and long-range interactions

Zeolite	Short-range parameters			
	A (10^3 kJ mol ⁻¹ Å ⁶)	Na	B (10^6 kJ mol ⁻¹ Å ¹²)	Na
Guest				
C	1.1759	0.5547	1.0169	0.6985
H	0.5802	0.1696	0.1118	0.0758
Long-range parameters				
Atom type	C	H	Si/Al	Na O
Charge	+0.07	-0.0175	+1.15	+1.0 -0.7

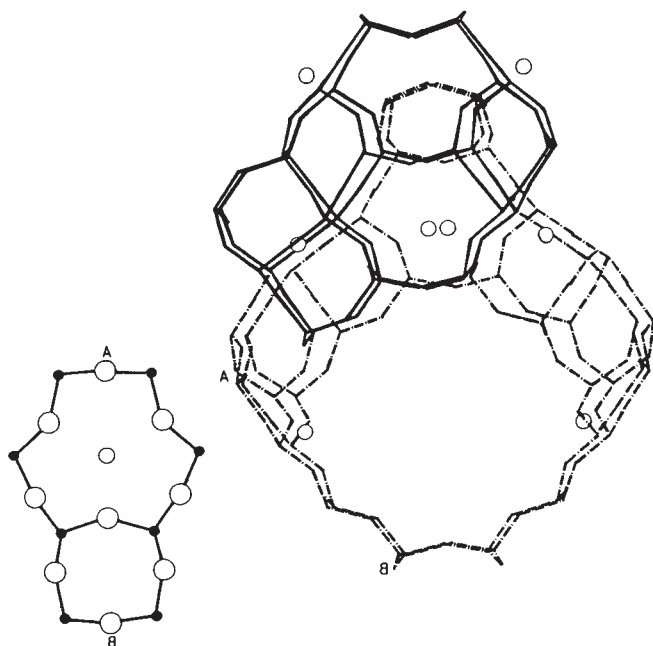


Fig. 1 Schematic drawing of part of the structure of zeolite X and Y showing the cation sites. The dashed lines show the super-cage inside which the methane is bound. Large circles, oxygen; small circles, sodium; filled circles, silicon/aluminium.

non-framework sites. In addition, many different cations and cation combinations of a binary or ternary nature can be accommodated⁶. At the same time, there is a paucity of reliable experimental information on the mobilities of hydrocarbons in zeolites, largely because the necessary measurements, such as variable (low) temperature solid-state nuclear magnetic resonance from which diffusion coefficients and activation energies can be obtained, have until recently been difficult or time-consuming^{7,8}. Computational approaches are therefore attractive alternatives and moreover, yield details at the microscopic level. Further, these approaches can be adapted to cope with the almost limitless variety of structural types that zeolites exhibit. In previous studies of methane sorbed in zeolites, a molecular mechanics procedure was used to obtain heats of adsorption and activation energies. Bezus *et al.*⁹ computed the thermodynamic properties of methane adsorbed in NaX with Si/Al=1.48; Nowak *et al.*¹⁰, following an earlier study¹¹ of pyridine in zeolite L, calculated the site and heat of adsorption for methane in NaY (Si/Al=3.0); and Fiedler and co-workers¹² have studied the influence of electrostatic field on the uptake of methane in faujasite and they have also reported a Monte Carlo treatment of hydrocarbons in type A zeolites^{13,14}. Higher hydrocarbons in faujasitic cages¹⁵ have been studied with regard to their thermodynamic properties and conformation as a function of temperature.

In this work we attempt to understand the behaviour of sorbed methane in terms of the temperature-dependence of its binding and mobility in the zeolite cage. Monte Carlo simulations¹⁶ used the Metropolis scheme in the canonical ensemble, at constant volume V and temperature T . The calculations were for a single methane corresponding to the zero coverage limit. Several simplifying assumptions were made for heuristic reasons: the zeolite framework, including the cations, is treated as rigid, as is methane. The only permitted degrees of freedom are the three translational and three rotational variables associated with the adsorbate. The interactions between methane and the atoms of the zeolitic framework were modelled in terms of a short-range Lennard-Jones (6-12) potential and a long-range coulombic term. Short-range interactions between the oxygen of Na-Y and

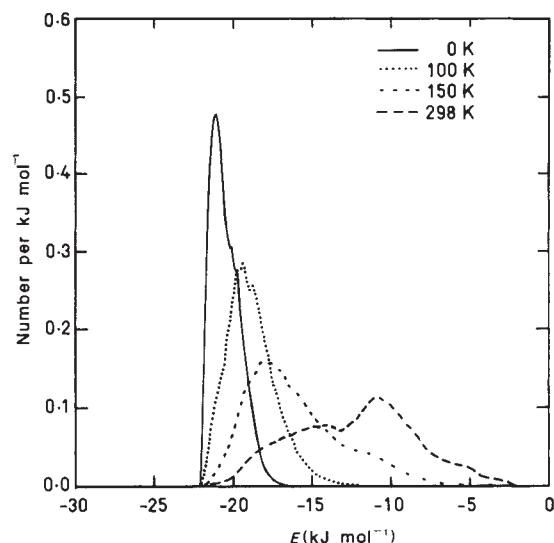


Fig. 2 Plots of the distribution functions of the potential energy of interaction of the methane with the zeolite NaY at different temperatures. Note the appearance of the shoulder on the higher-energy side around 170 K.

the atoms in methane only were considered, as close approach of methane to the Si or Al atoms is not feasible. No distinction was made between the Si and Al. Table 1 gives parameters used for the calculation of short-range and long-range interactions, and the charges were derived from molecular orbital calculations by the MNDO (modified neglect of diatomic overlap) approach¹⁷. Summations were taken up to a cut-off radius of 10 Å on a single unit cell of the zeolite and one methane molecule. Each Monte Carlo step consisted of a random displacement of centre of mass and a random rotation through a randomly selected axis passing through the centre of mass. Typically, equilibration was over $(20-50) \times 10^3$ steps and averaging over another $(200-300) \times 10^3$ steps. All calculations were carried out for a Si/Al=3.0 with $a=24.85$ Å (ref. 18). For this ratio, monovalent cations occupy sites SI and SII completely (Fig. 1).

Table 2 gives differential internal energy changes and isosteric heats of adsorption calculated by the Monte Carlo method for different temperatures. These values agree with the experimental ones¹⁹⁻²², which vary between -14.9 and -16.5 kJ mol⁻¹ for methane adsorbed in Na-X at $T=323$ K, and with those values calculated by Kiselev *et al.*¹⁹ (15.5 kJ mol⁻¹) for methane in Na-Y at 323 K. Minimization of the energy with respect to the position and orientation of methane led to a site near the central four-membered ring containing the hexagonal rings of the super-cage. At this position the C of methane is nearly equidistant (~ 5 Å) from two of four cations at S II sites. The energy at this site is -22.1 kJ mol⁻¹.

Figure 2 gives the distribution of energy of interaction of methane with the zeolite at different temperatures. At 50 K, the potential-energy distribution function (PEDF) is narrow, sharp

Table 2 Computed differential internal energy change ($-\Delta U$) and isosteric heat of adsorption (q_{st}) of methane in Na-Y

T (K)	$-\Delta U$ (kJ mol ⁻¹)	q_{st} (kJ mol ⁻¹)
10	21.8	21.9
50	20.45	20.85
100	18.8	19.6
150	16.88	18.08
170	15.2	17.1
220	13.9	15.2
298	11.3	13.8

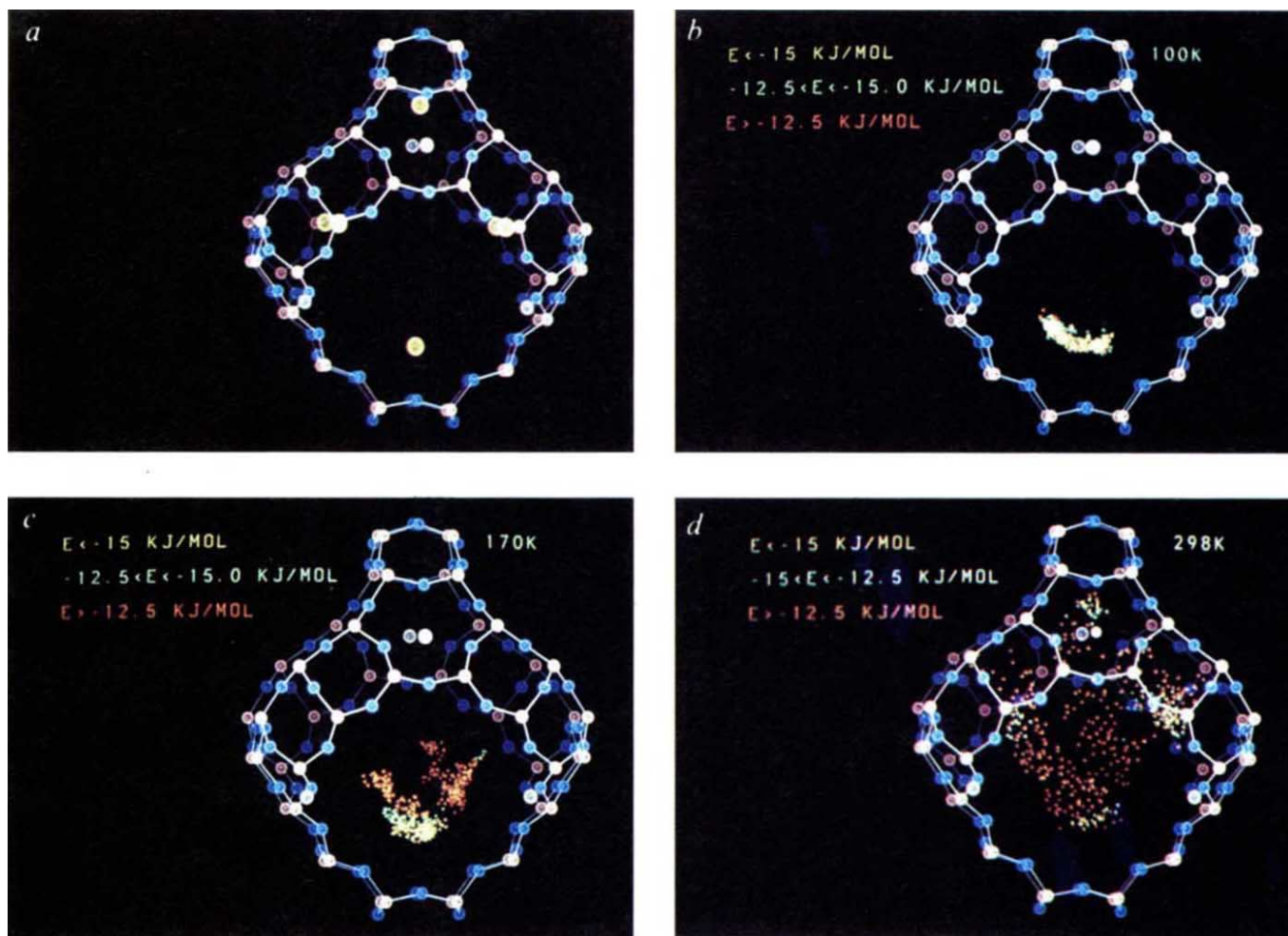


Fig. 3 *a*, Minimum energy sites occupied by methane within the supercage are shown in yellow. The positions occupied by methane at 100, 170 and 298 K over a 1,000 Monte Carlo steps are shown in *b*, *c* and *d*, respectively, colour-coded as indicated according to the potential interaction energy. The increased mobility of methane when its interaction energy corresponds to the shoulder in Fig. 2 is apparent.

and unimodal with the peak maximum at $\sim -20.5 \text{ kJ mol}^{-1}$, confirming that the molecules are localized at low temperatures. The asymmetric nature of the peak in the PEDF is in contrast to the symmetric shape usually found for liquids¹⁶. It arises from the existence of a lower bound ($-22.1 \text{ kJ mol}^{-1}$) for the interaction energy of methane with the zeolite. At significantly high temperatures the distribution functions should become symmetric as the minimum energy site become increasingly less occupied. At 100 K, the peak maximum has shifted to $-19.5 \text{ kJ mol}^{-1}$ and is still unimodal, but by 150 K, a small tail appears on the higher energy side of the PEDF (not shown in Fig. 2). This tail in turn develops into a shoulder by 170 K, and at 200 K the PEDF is clearly bimodal, with the maximum at $\sim -17.5 \text{ kJ mol}^{-1}$ and distinct shoulder at $\sim -12 \text{ kJ mol}^{-1}$. The shoulder grows more intense with increasing temperature, and by 298 K it develops into a main peak while the low-energy one at $\sim -18 \text{ kJ mol}^{-1}$ becomes a shoulder.

Figure 3 shows the positions occupied by methane over 1,000 Monte Carlo steps at three different temperatures: 100, 170 and 298 K. At 100 K, the molecule is essentially confined to a small region corresponding to the minimum energy in the supercage. This position is associated with the sharp peak in the PEDF at this temperature, and as a consequence of the symmetry, there are six equivalent sites of this type in the supercage. At 170 K, methane hops from one minimum energy site to another, and the higher-energy peak in the PEDF corresponds to the regions in the supercage that are occupied during the process of hopping. The molecule appears to be more mobile when it is occupying

these higher-energy sites, which it does for $\sim 16\%$ of the time at 170 K. At 220 K, 35% of its time is spent in such positions. The 298-K distribution shown in Fig. 3 reveals that the molecule is now highly mobile, spending rather little time at the minimum energy positions.

Heat-capacity measurements for methane and ethane in zeolite 5A by Parsonage and co-workers^{23,24} show characteristic peaks in the temperature range 10–350 K. The temperature-dependence of the heat capacity can be explained in terms of the Schottky model comprising potential wells of different depths. Based on this, it has been suggested that the difference between the potential minima near the surface of the cavity are in the range $4\text{--}9 \text{ kJ mol}^{-1}$. The potential minima near the surface of the cavity is probably determined by the local features, namely the cations and the 4- and 6-rings. Given that the structures of A and Y have much in common, similar values for the difference between the potential minima are to be expected for zeolite Y. The values obtained by us (between ~ 5.5 and $\sim 7.5 \text{ kJ mol}^{-1}$) seem to support this. Infrared²⁵ and neutron-scattering²⁶ measurements of methane in A found that at $\sim 200 \text{ K}$, methane is bound close to the cavity wall corresponding to the potential minima, and on increasing the temperature to 300 K, methane was completely delocalized. Thus, the results obtained by us are in good agreement with the experiment.

Earlier simulations of methane in zeolite are approximate in that the methane or guest species is represented by a single interaction site, and only short range interactions are considered between methane and the zeolite cavity. This study uses a

five-site model for methane and takes into account the long-range coulombic interactions between methane and the zeolite cavity. We therefore believe that these results represent a more accurate and realistic description of the energetics and spatial distribution as a function of temperature.

It is clear that, for small molecules such as methane, a wide range of positions will be sampled in a catalyst held at its working temperature, and the minimum energy sites need not have a significant role in reactions. Calculations of this type should therefore be of value in future attempts to model catalysis under realistic conditions.

We are grateful to Professor N. G. Parsonage for helpful comments. We acknowledge support from the Science and Engineering Research Council for part of this work.

Received 21 August; accepted 5 November 1987.

1. Nicholson, D. & Parsonage, N. G. *Computer Simulation and the Statistical Mechanics of Adsorption* (Academic Press, London, 1982).
2. Abraham, F. F. *Adv. Phys.* **35**, 1-111 (1986).
3. Ramdas, S., Thomas, J. M., Klinowski, J., Fyfe, C. A. & Hartman, J. S. *Nature* **292**, 228-230 (1981).

4. Klinowski, J., Ramdas, S. & Thomas, J. M. *JCS Faraday Trans. II* **78**, 1025-1050 (1982).
5. Newsam, J. M. *Science* **231**, 1093-1096 (1986).
6. Mortier, W. J. *Compilation of Extra Framework Sites in Zeolites* (Butterworths, London, 1982).
7. Kärger, J. & Pfeifer, H. *Zeolites* **7**, 90-107 (1987).
8. Bülow, M., Struve, P. & Mietk, W. *Z. phys. Chem.* **267**, 613-616 (1986).
9. Bezus, A. G., Kiselev, A. V., Lopatkin, A. A. & Du, P. Q. *JCS Faraday Trans. II* **74**, 367-376 (1978).
10. Nowak, A. K. & Cheetham, A. K. in *Proc. 7th Int. Zeolite Conf.* (eds Murakami, Y., Lijima, A. & Ward, J. W.) (Kodansha, Elsevier, 1986).
11. Wright, P. A., Thomas, J. M., Cheetham, A. K. & Nowak, A. K. *Nature* **318**, 611-615 (1985).
12. Fiedler, K. *et al.* in *Proc. 5th Int. Zeolites Conf.* (ed. Rees, L. V. C.) (Heyden, London, 1980).
13. Kretschmer, R. G. & Fielder, K. *Z. phys. Chem.* **258**, 1045-1058 (1977).
14. Fiedler, K. & Stach, H. *Chem. Tech. (Leipzig)* **36**, 104-108 (1984).
15. Fiedler, K., Stach, H. & Schirmer, W. in *Theorie der Adsorption an Zeolithen* (Sitzungsberichte der AdW der DDR) **17/N**, 5-40 (1982).
16. Yashonath, S. & Rao, C. N. R. *Proc. R. Soc. Lond. A* **400**, 61-70 (1985).
17. Dewar, M. J. S. & Thiel, W. J. *Am. Chem. Soc.* **99**, 4899-4907 (1977).
18. Fitch, A. N., Jobic, H. & Renouprez, A. *J. phys. Chem.* **90**, 1311-1315 (1986).
19. Kiselev, A. V. & Du, P. Q. *JCS Faraday Trans. II* **77**, 17-23 (1981).
20. Barrer, R. M. & Sutherland, J. W. *Proc. R. Soc. Lond. A* **237**, 439-449 (1956).
21. Habgood, H. W. *Can. J. Chem.* **12**, 2340-2346 (1964).
22. Neddenriep, R. J. *J. Colloid Interface Sci.* **28**, 293-299 (1968).
23. Stroud, H. J. F., Richards, E. & Parsonage, N. G. *JCS Faraday Trans. I* **72**, 942-954 (1976).
24. Richards, E., Stroud, H. J. F. & Parsonage, N. G. *JCS Faraday Trans. I* **72**, 1759-1769 (1976).
25. Cohen de Lara, E. & Seloudoux, R. *JCS Faraday Trans. I* **79**, 2271-2276 (1983).
26. Cohen de Lara, E. & Kahn, R. *J. Phys., Paris* **42**, 1029-1038 (1981).

Fractionation of nitrogen isotopes in a synthetic diamond of mixed crystal habit

S. R. Boyd*, C. T. Pillinger*, H. J. Milledge†, M. J. Mendelsohn† & M. Seal‡

* Planetary Sciences Unit, Department of Earth Sciences, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK
 † Crystallography Unit, Department of Geology, University College, London, UK
 ‡ D. Drukker and Zn. N.V., Sarphatikade 12, Amsterdam, The Netherlands

Nitrogen is the most abundant trace element in natural¹ and meteoritic² diamond and finds its way into synthetic diamonds during their manufacture³. In the first study of the isotopic composition in a synthetic diamond, we have found that nitrogen incorporated in a large (~0.5 cm diameter) cubo-octahedron is isotopically fractionated between the zones of different crystal habit. The cubic {100} sector of the stone, which contains about half the concentration of nitrogen, is enriched by ~4.5% in the heavy (¹⁵N) isotope of the element relative to the octahedral {111} growth region. This is the largest isotopic fractionation yet observed that can be attributed to the influence of crystal structure and demonstrates again the idiosyncratic behaviour of diamond with respect to nitrogen.

Nitrogen in natural diamond has been extensively studied by a variety of spectroscopic techniques so that it is known to exist as single atoms, pairs (A form), groups of three (N3 centres), four or more (B form) and thousands to millions (present in platelets) of atoms⁴. Measurement of its isotopic composition as δ values, where

$$\delta(\%) = \left[\frac{(^{15}\text{N}/^{14}\text{N})_{\text{sample}} - (^{15}\text{N}/^{14}\text{N})_{\text{standard}}}{(^{15}\text{N}/^{14}\text{N})_{\text{standard}}} \right] \times 1,000$$

can be informative in the recognition of the growth history of kimberlitic diamonds⁵. In meteoritic diamonds, thought to have originated before the formation of the Solar System⁶, $\delta^{15}\text{N}$ values are highly anomalous². Because nitrogen isotopic compositions are being considered as provenance related, we investigated the scale of nitrogen isotopic fractionation during diamond crystallization. To obtain relevant information, we have determined the spatial distribution of nitrogen and its isotopic abundance in a large (~0.5 cm diameter) synthetic diamond in which cubic and octahedral growth sectors could be distinguished by cathodoluminescence⁷. Isotope measure-

ments were made on ions of $m/z = 28$ and 29 using high sensitivity static mass spectrometry, on nitrogen gas produced by combustion of fragments (250 × 250 μm) laser sectioned^{5,8} from a 400- μm thick slice cut parallel to the {110} orientation. The instrument, built in our laboratory, the measurement technique and an assessment of the overall performance, including steps taken to avoid interference from isobaric species, are described elsewhere^{5,9,10}. In some instances multiple fragments were taken to provide enough gas for mass spectrometric determination; combining fragments was considered permissible on the basis of matching Fourier transform infrared (FTIR) spectra^{4,5}.

Figure 1 shows nitrogen abundance, $\delta^{15}\text{N}$ and $\delta^{13}\text{C}$ values (determined conventionally) for various sections of the diamond plate. The octahedral {111} portion of the diamond, which is yellow, contains twice as much nitrogen as the white cubic {100} part as expected from previous studies⁷ and in agreement with FTIR data suggesting differences in nitrogen distribution for this stone. The nitrogen-rich areas have $\delta^{15}\text{N}$ values close to atmospheric (defined as 0‰), decreasing (-0.4 to -2.5‰) from the centre outwards, but the nitrogen-poor cubic regions have $\delta^{15}\text{N}$ values that increase from +30.7‰ to a maximum of +42.1‰ towards the edge of the plate. All $\delta^{15}\text{N}$ values were measured to a precision of $\pm 0.5\%$ (1 σ). In the brightly luminescing band between the sectors, which outcrops on {113}, and at the extreme edge of the cubic zone, FTIR suggests extremely low nitrogen abundance; no mass spectrometric measurements were made on such fragments. As well as being intrinsically interesting it is important to understand the enrichment of ¹⁵N because of possible effects on physical properties, notably thermal conductivity¹¹. On a lesser scale there could be some structure in the $\delta^{13}\text{C}$ measurements in that ¹³C abundance is higher at the centre than the edge (-24.09 to -24.69‰ octahedral; -23.86 to -24.40‰ cubic). The $\delta^{13}\text{C}$ values reported here have all been determined to a precision of $< \pm 0.03\%$ so the growth trends they reveal from interior to exterior may be real. The single exception to the above trends is the fragment close to the edge of the octahedral zone ($\delta^{13}\text{C} = -21.45$; 130 p.p.m. N, $\delta^{15}\text{N} = -0.6\%$), which seems anomalous in respect of both elements. It is not clear if there is a difference in $\delta^{13}\text{C}$ between cubic and octahedral diamond because of inappropriate sampling.

To understand the isotopic systematics of the diamond it would be useful to know the conditions under which it was grown, the nature of the materials used in its production and their isotopic composition. However, the stone we studied came from de Beers' Industrial Diamond Corporation and its previous history is undocumented. It is possible, however, to discuss the methods for producing synthetic diamonds in general terms.

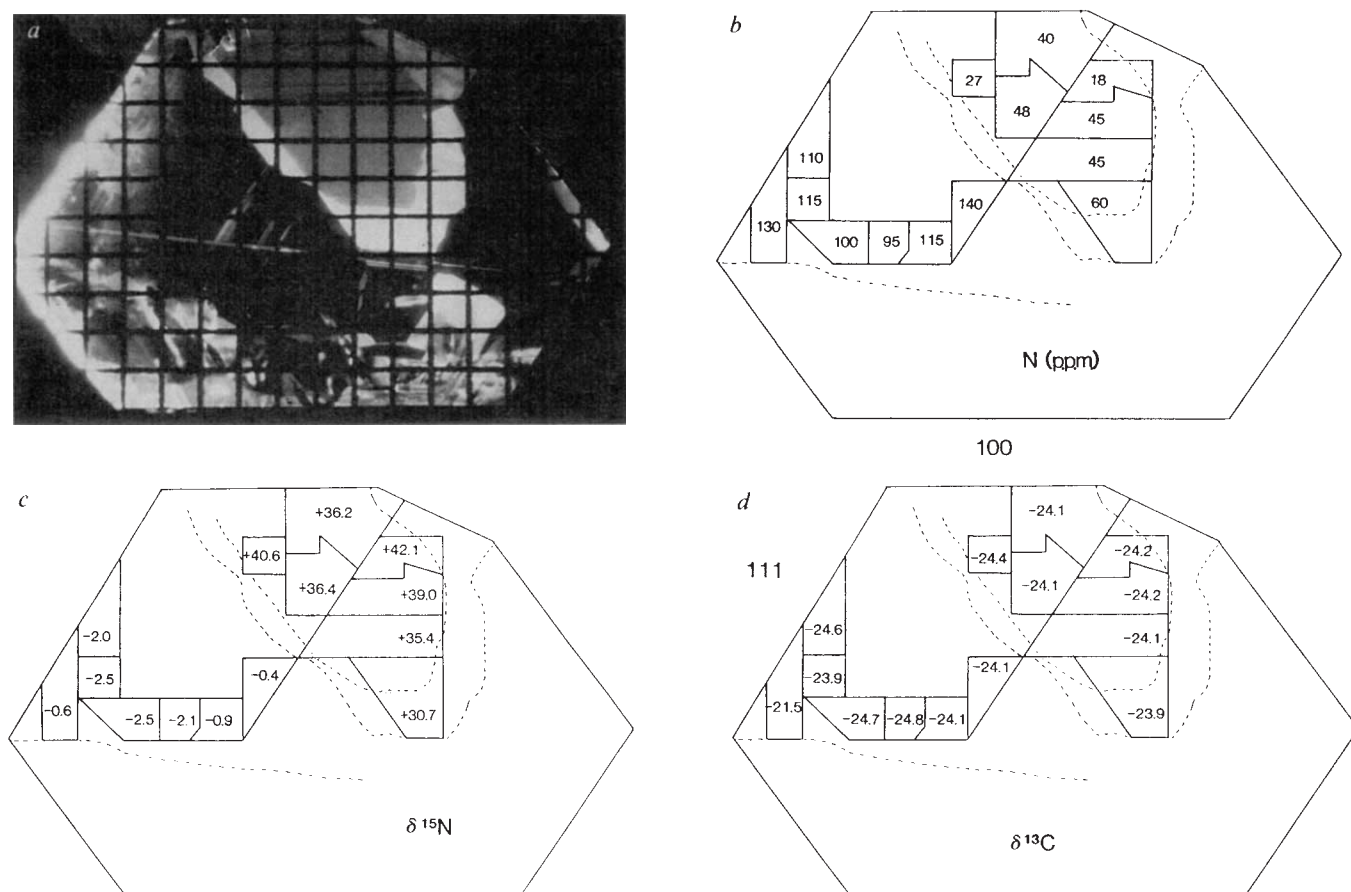


Fig. 1 *a*, A print from a coloured cathodoluminescence picture of the synthetic diamond plate used in the experiments. Here brightly luminescing areas are dark. The 250 $\mu\text{m} \times 250 \mu\text{m}$ grid on the diamond has been cut with a Nd-YAG laser so that selected areas may be broken out of the specimen. Nitrogen abundance (*b*), $\delta^{15}\text{N}$ (*c*) and $\delta^{13}\text{C}$ (*d*) for the cubic ({100}) and octahedral ({110}) sector are shown in appropriate accompanying diagrams.

Most small synthetic diamonds (<1 mm in size) are grown rapidly (during a few minutes) from graphite in the presence of solvent metals such as cobalt, iron or nickel¹² at temperatures ~1,720 K and pressures ~55–60 kbar, although there are other methods¹³. Large diamonds such as that described here are obtained over a period of days using a reconstitution technique¹⁴, where small diamonds, almost certainly synthetics, act as the carbon source to avoid volume changes during growth. In this method, diamond is dissolved in a hot source region, typically 1,720 K, and reprecipitated in a cooler seeded region where the temperature differential preferred for relatively slow growth (2.5 mg h⁻¹)³ is ~30 K. The driving force for the process is the solubility gradient between the two regions. The morphology of the reconstituted diamond is determined by the exact thermodynamic conditions selected. Thus octahedral diamond is produced at high temperature in a pressure regime close to the Berman–Simon graphite–diamond equilibrium curve¹⁵. Cubic diamond normally grows at lower temperatures and pressures well into the diamond stability field; cubo-octahedra are formed at intermediate pressures and temperatures. Minor amounts of {110} and {113} rhombohedral growth of diamond can also be found in both large and small synthetic stones⁷ and this morphology can be enhanced by adding water to the system¹⁶. The presence of nitrogen appears to promote growth, but too much added as azide, cyanide etc., has deleterious effect¹⁴. It is more common to decrease the residual nitrogen content of synthetic diamonds in experiments to obtain gems, heat-sinks or semiconductors doped with boron, by introducing a getter metal such as Al, Ti or Zr¹⁴.

Despite the lack of knowledge about the generation of the diamond studied here we can rationalize its isotopic

heterogeneity against possible growth models. One previous attempt to measure the carbon isotopic fractionation during the conversion of graphite to diamond was made by Hoering¹⁷, who found an enrichment of 0.3%, but this value was within the uncertainty of his technique¹⁸. Our observations suggest that ¹³C-enriched carbon is the first to crystallize in both cubic and octahedral zones, which is consistent with a process whereby nickel metal solvent becomes supersaturated with carbon before precipitation occurs rather than decomposition of a carbide¹² or a diffusion-controlled operation. Crystal structure can influence isotopic properties¹⁹; if there is a difference in vibrational frequency, manifested as a change in density between polymorphs, the more closely packed well-ordered structure tends to concentrate the heavy isotope. As there is no density difference between cubic and octahedral diamond (unlike diamond and graphite), and in any case because fractionation factors should be small at high temperatures, there is no reason to expect a change in isotopic composition in going from one growth morphology to another. The $\delta^{13}\text{C}$ values acquired here do not contradict such assumptions, although the octahedral sector does exhibit slightly lower ¹³C abundance. Future studies will employ a laser technique that excises separate growth sectors parallel to the crystal faces and dissects them to sample distinct growth horizons.

We turn now to the nitrogen abundances and isotopic compositions: it is known that nitrogen is incorporated into the cubic and octahedral sectors of synthetic diamonds to different extents, although no isotopic measurements have previously been attempted. Strong and Chrenko³ suggested octahedral zones with up to 100 p.p.m. nitrogen but other regions only 5–7 p.p.m. This is an even more extreme partition than observed here and

would immediately imply that conditions were favourable for an equilibrium isotopic fractionation effect. The trends in $\delta^{15}\text{N}$ are, at face value, entirely consistent with an equilibrium process with the octahedral and cubic zones showing contrasting preferences for the light and heavy isotopes, respectively. Since $\Delta^{15}\text{N}$ ($\delta^{15}\text{N}_{\text{cubic}} - \delta^{15}\text{N}_{\text{octahedral}}$) was increasing as the stone grew, possibly some condition, most likely temperature, in the seed region was changing so that the fractionation factor became larger. Alternatively a limited nitrogen source was being depleted during growth. The fact that the fractionations are large and occurred at high temperature, however, point to kinetic control or lack of attainment of equilibrium.

Not knowing the source of the nitrogen, we could not explore equilibrium fractionation more fully. If the starting nitrogen had a $\delta^{15}\text{N}$ close to atmospheric, say $0 \pm 2.5\%$, which is likely, then only the cubic face has any marked fractionation effect; any other starting $\delta^{15}\text{N}$ leads to more complicated interpretations. Originally it was unknown how impurity nitrogen was accommodated in the diamond structure—interstitial or substitutional, singly or paired¹—but it is now established that initially, in both natural and synthetic diamonds, nitrogen is incorporated as single atoms at substitutional sites^{20–22}. Indeed, the yellow colour of the octahedral sector of our diamond confirms this interpretation. To give single atoms, molecular nitrogen must dissociate, and to insert a nitrogen into diamond distorts the lattice considerably by moving away from the tetrahedral carbon atoms to lengthen the bonds by $\sim 10\%$. As a consequence of both processes, there are clear opportunities for isotopic effects. Synthetic diamonds appear to form in layers a few micrometres thick, with nucleation and growth rate controlled by a delicate balance between the supply of carbon to the edge of the advancing step, the adsorption of impurities on the old surface and their diffusion away from the base of the step¹⁴. There are several interdependent but unconstrained variables here that prevent us from speculating on the likely repercussions for isotopic fractionation.

As an alternative to the above explanation of the $\delta^{15}\text{N}$ measurements, it is conceivable that if a gettering process was used during manufacture of our stone to remove unwanted nitrogen from the system, it could be in some way implicated. However, this would require gettering to be operative after crystallization of the diamond so that isotopic enrichment in the cubic sector could be achieved by preferential diffusion of light atoms. Differences in abundances and $\delta^{15}\text{N}$ would then be observable during mass spectrometric measurement if the metal nitride formed by gettering survived the combustion reaction used to liberate nitrogen from the diamond, that is, the intervention of an experimental artefact. Against such an interpretation is the observation, based on infrared analysis, that nitrogen does not exist in the synthetic diamond in the paired form; it appears to have had a restricted mobility at the temperature of synthesis²³. Grady *et al.*²⁴ have reported that osbornite (TiN) will combust to release nitrogen at $\sim 1,273\text{ K}$. Future studies can greatly clarify this discussion because the distribution of metal inclusions and their composition can be established before destruction of the diamond by combustion. Our mass spectrometric techniques may even be able to measure the abundance and isotopic composition of any nitrogen associated with metal.

The $\delta^{15}\text{N}$ variations observed here between cubic and octahedral diamond are the largest yet encountered for a structurally controlled process. Given that octahedral and cubic growth diamond have exactly the same arrangement of atoms, it is most surprising that isotopic fractionation is occurring between different regions of the same crystal. Even when a compound such as calcium carbonate exists in entirely different crystal systems, such as rhombohedral (calcite) and orthorhombic (aragonite) the $\delta^{13}\text{C}$ change at 298 K is $1.8\text{--}4.0\%$ ^{25,26} and the $\delta^{18}\text{O}$ shift (0.6%) is even less^{26–28}. One of the largest effects known is between diamond and graphite calculated to be 11.5%

at 273 K , decreasing to 0.4% at $1,273\text{ K}$ (ref. 20). Similarly another calculated, rather than measured, $\delta^{18}\text{O}$ difference is that estimated for α and β quartz as $0.8\text{--}1.5\%$ at the transition temperature, 846 K (refs 29, 30). Perhaps not surprisingly, the biggest previously observed^{31,32} fractionations accompanied either a change of state (ice/water at 273 K , $\delta^{18}\text{O} = 3.0$ and $\delta\text{D} > 19.5\%$) or the involvement of hydrogen ($\text{H}_2\text{O}/\text{D}_2\text{O}$ at 298 K , $\delta^{18}\text{O} > 15\%$). The obvious difference between the systems investigated and nitrogen in diamond is that until now major, rather than trace, elements have been considered. It could be profitable to evaluate other situations, for instance the crystallographic changes in the structure of ice, from amorphous to cubic to hexagonal as its temperature is changed, particularly as ice traps gaseous species as clathrates³³. Diamonds produced by chemical vapour deposition in a microwave discharge³⁴ could also be interesting, since $\delta^{15}\text{N}$ values up to 423% have been encountered in plasma-type environments^{35,36}.

The fractionation of nitrogen isotopes between crystal forms is relevant to processes involving natural diamonds. We have already examined⁵ nitrogen isotopes in coated stones where material of cubic habit has developed on an octahedral core, that is, where growth is sequential rather than simultaneous. Octahedral cores from a series of diamonds believed to have a common provenance had variable $\delta^{13}\text{C}$ (-5.0 to -9.8%) and $\delta^{15}\text{N}$ (-3.4 to $+13.4\%$), whereas cubic coats and individual cubes were characterized by more restricted values ($-6.5\text{--}7.5$ and -4.0 to -8.0% respectively). These observations are in the opposite direction to what we have seen in the synthetic stone. However, what could be important in the natural case is the fractionation between the initial reservoir and the crystalline form of diamond measured. There is no way yet of knowing the reservoir composition, but the possible availability of large-scale fractionation mechanisms for nitrogen, capable of working at high temperature and pressure in the deep Earth, might have implications for core-mantle processes. Among the more appropriate types of natural diamond to investigate would be the unusual specimens that show a central cross after etching^{37,38} or viewed by cathodoluminescence³⁹ to provide evidence of simultaneous growth of mixed habit. If the fractionation observed for synthetics is reproduced, then understanding its magnitude offers the possibility of determining the conditions in the source region. It is interesting that cliftonite, thought at one time to be a pseudomorph after diamond, and lonsdaleite⁴⁰, both found in iron meteorites, also exhibit sectorized zoning of cubo-octahedral habit. Sectorized zoning is not uncommon in other minerals, having been seen in, for instance, staurolite⁴¹, pyroxenes⁴² and carbonates⁴³.

From studies of included silicate and other minerals, natural diamond is known to grow at similar temperature and pressure as synthetic stones⁴⁴, although other conditions such as the nature of the melt and oxygen fugacity are clearly different. After incorporation of nitrogen as single atoms, mass-dependent aggregation processes, due to prolonged residence in high-temperature regions within the Earth²¹, may overprint isotopic signatures introduced by crystallographic preferences. The nitrogen isotopic systematics of diamond, both natural and synthetic, remain far from understood but can be addressed by the advanced techniques now available.

We thank Drs I. P. Wright and M. H. Thiemens for their constructive comments on this manuscript. S.R.B. thanks NERC for a research studentship.

Received 7 July; accepted 26 November 1987.

1. Kaiser, W. & Bond, W. L. *Phys. Rev.* **115**, 857–863 (1959).
2. Lewis, R. S., Anders, E., Norris, S. J., Wright, I. P. & Pillinger, C. T. *Nature* **305**, 767–771 (1983).
3. Strong, H. M. & Chrenko, R. M. *J. phys. Chem.* **75**, 1838–1843 (1971).
4. Davies, G. *Diamond* (Adam Hilger, Bristol, 1984).
5. Boyd, S. R. *et al. Earth planet. Sci. Lett.* **86**, 343–351 (1987).
6. Lewis, R. S., Ming, T., Wacker, J. F., Anders, E. & Steel, E. *Nature* **326**, 160–162 (1987).
7. Woods, G. S. & Lang, A. R. *J. Crystal Growth* **28**, 215–226 (1975).
8. Swart, P. K., Pillinger, C. T., Milledge, H. J. & Seal, M. *Nature* **303**, 793–795 (1983).

9. Wright, I. P., Boyd, S. R., Franchi, I. A. & Pillinger, C. T. *J. phys. E* (submitted).
10. Exley, R. A., Boyd, S. R., Mathey, D. P. & Pillinger, C. T. *Earth planet. Sci. Lett.* **81**, 163-174 (1987).
11. Goodman, C. H. L. *J. Crystal Growth* **75**, 1-14 (1986).
12. Giardini, A. A. & Tydings, J. E. *Am. Miner.* **47**, 1393-1421 (1962).
13. Wedlake, R. J. in *Properties of Diamond* (ed. Field, J. E.) 501-535 (Academic, London, 1970).
14. Wentdorf, R. H. *J. phys. Chem.* **75**, 1833-1837 (1971).
15. Berman, R. & Simon, F. Z. *Elektrochem.* **59**, 333-345 (1955).
16. Kanda, H., Setaka, N., Ohsawa, T. & Fukunaga, O. *J. Crystal Growth* **60**, 441-444 (1982).
17. Hoering, T. C. *Carnegie Inst. Wash. Pap. Geophys. Lab. no. 64* (1961).
18. Bottinga, Y. *Earth planet. Sci. Lett.* **5**, 301-307 (1969).
19. O'Neil, J. R. *Rev. Mineral.* **16**, 1-40 (1986).
20. Harris, J. W., Hawthorne, J. B., Oosterveld, M. M. & Wehmeyer, E. *Phys. Chem. Earth* **9**, 765-783 (1973).
21. Evans, T. *Contemp. Phys.* **17**, 45-70 (1976).
22. Chrenko, R. M., Tuft, R. E. & Strong, H. M. *Nature* **270**, 141-144 (1977).
23. Evans, T. & Qi, Z. *Proc. R. Soc. A* **381**, 159-178 (1982).
24. Grady, M. M., Wright, I. P., Carr, L. P. & Pillinger, C. T. *Geochim. cosmochim. Acta* **50**, 2799-2813 (1986).
25. Robinson, M. & Clayton, R. N. *Geochim. cosmochim. Acta* **33**, 997-1002 (1969).
26. Golyshev, S. I., Padalko, N. L. & Pechenkin, S. A. *Geochem. Int.* **10**, 85-99 (1981).
27. Tarutani, T., Clayton, R. N. & Mayeda, T. K. *Geochim. cosmochim. Acta* **33**, 987-996 (1969).
28. O'Neil, J. R., Clayton, R. N. & Mayeda, T. K. *J. chem. Phys.* **51**, 5547-5558 (1969).
29. Shiro, Y. & Sakai, H. *Bull. chem. Soc. Japan* **45**, 2355-2359 (1972).
30. Kawabe, I. *Geochim. cosmochim. Acta* **42**, 613-621 (1978).
31. O'Neil, J. R. *J. phys. Chem.* **72**, 3683-3684 (1968).
32. Staschewski, D. *Ber. Bunsenges. phys. Chem.* **69**, 426-440 (1965).
33. Bar-Nun, A., Herman, G., Laufer, D. & Rappaport, M. C. *Icarus* **63**, 317-332 (1985).
34. Fukunaga, K., Matsuda, J., Nagao, K., Miyamoto, M. & Ito, K. *Abstr. Twelfth Symp. Antarctic Meteorites* 116-118 (1987).
35. Manuccia, T. J. & Clark, M. D. *Appl. Phys. Lett.* **28**, 372-374 (1976).
36. Arrhenius, G., Fitzgerald R., Markus, S. & Simpson, C. *Astrophys. Space Sci.* **55**, 285-297 (1978).
37. Seal, M. *Nature* **212**, 1528-1531 (1965).
38. Harrison, E. R. & Tolansky, S. *Proc. R. Soc. A* **279**, 490-496 (1964).
39. Lang, A. R. in *Properties of Diamond* (ed. Field, J. E.) 425-469 (Academic, London, 1979).
40. Brett, R. & Higgins, G. T. *Geochim. cosmochim. Acta* **33**, 1473-1484 (1969).
41. Hollister, L. S. *Am. Miner.* **55**, 742-766 (1970).
42. Carpenter, M. A. *Am. Miner.* **65**, 313-321 (1980).
43. Reeder, R. J. & Prosky, J. L. *J. Sedim. Petrol.* **56**, 237-247 (1986).
44. Meyer, H. O. A. & Tsai, H.-M. *Miner. Sci. Engng* **8**, 242-261 (1976).

Atmospheric wet sulphate deposition and lakewater chemistry

T. J. Sullivan, J. M. Eilers, M. R. Church, D. J. Blick, K. N. Eshleman, D. H. Landers & M. S. DeHaan

Environmental Research Laboratory—Corvallis, 200 SW 35th Street, Corvallis, Oregon 97333, USA

Atmospheric sulphur deposition and its impacts on surface-water chemistry and biota have been extensively studied in several areas of the United States and elsewhere during the past two decades^{1,2}. In the absence of regionally based statistical (probability) sampling, however, it has not been possible to extrapolate the results from intensive small-scale studies to a regional population. Here we present data from the first regional probability sampling of lakes in the United States^{3,4}, indicating that median lake sulphate concentrations correlate highly with estimated wet sulphate deposition at lake sites. The data further suggest that acid deposition has depleted acid-neutralizing capacity (ANC) in low-base-cation lakes, and that some high-base-cation lakes receive substantial sulphate from internal watershed sources. Data suggest that acidic lakes greater than 4 hectares in surface area are generally not found in areas receiving low sulphate deposition, whereas most acidic lakes in the northeastern United States are currently acidic because of high sulphate relative to base cation concentrations.

Lake chemistry data were taken from the recent National Lake Survey (NLS)^{3,4}, conducted by the US Environmental Protection Agency (EPA). Details on statistical design, sampling, analytical methodologies, and quality assurance protocols were provided by Linthurst *et al.*³ and Landers *et al.*⁴. Estimates of wet SO₄²⁻ deposition at each lake site were derived from interpolated precipitation volume and chemistry data. For the northeastern USA, precipitation volumes were obtained from 30-year precipitation norm values, 1951-80, for 500 stations (National Climate Center, NOAA), and precipitation chemistry data were derived from NADP/NTN estimates for the 1-year period before lake sampling. In the upper midwestern and southeastern USA, interpolations based on precipitation volumes were not available for the long term. We therefore used interpolations from 3 years of precipitation volume and chemistry data from 162 stations in the Acid Deposition System, 1980-82 (ref. 5). Deposition estimates were assigned by interpolating both SO₄²⁻ concentration and precipitation volumes. Because of the sparse monitoring of precipitation chemistry and volume in the western United States, an SO₄²⁻ deposition value was derived from 1981-84 annual averages for one or more NADP/NTN sites considered most appropriate for the lake study areas, rather than interpolating to each site⁶. Lake chemistry data in the Northeast were corrected for estimated

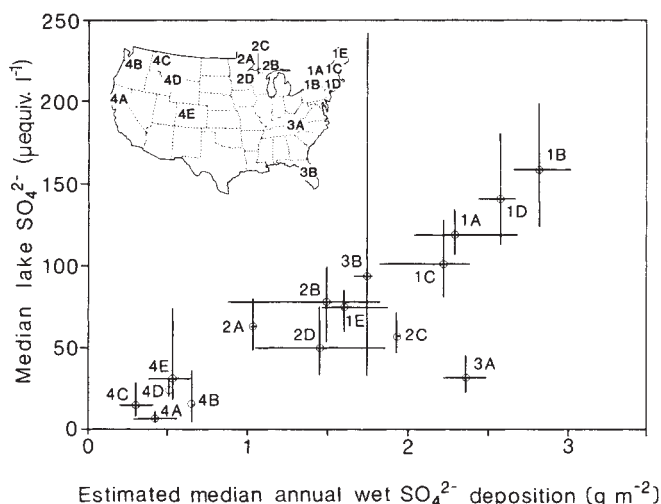


Fig. 1 Median (circle) and 25th and 75th percentiles (bars) of measured lake [SO₄²⁻] as a function of wet SO₄²⁻ deposition for subregions of the National Lake Survey. Deposition was estimated at each lake site for Eastern and Upper Midwestern lakes from interpolated precipitation volume and chemistry data. Estimated wet SO₄²⁻ deposition in the Western Region was taken from Hidy and Young⁶. Map insert depicts regions and subregions included in the EPA National Lake Survey. Subregion designations are as follows: 1A, Adirondacks; 1B, Poconos/Catskills; 1C, Central New England; 1D, Southern New England; 1E, Maine; 2A, Northeastern Minnesota; 2B, Upper Peninsula of Michigan; 2C, North-central Wisconsin; 2D, Upper Great Lakes Area; 3A, Southern Blue Ridge; 3B, Florida; 4A, California; 4B, Pacific Northwest; 4C, Northern Rockies; 4D, Central Rockies; 4E, Southern Rockies. SO₄²⁻ values were not corrected for marine contributions.

marine contributions of SO₄²⁻, Cl⁻, and base cations and estimated anthropogenic NaCl inputs. Marine Cl⁻ was estimated by using an empirical relation between Cl⁻ and distance to the coast from a data set of 'undisturbed' lakes, defined as those without paved roads in the watersheds based on examination of 1:250,000-scale US Geological Survey topographic maps. These estimated values for marine Cl⁻ were used for sea-salt corrections by assuming that the proportion of ions in the precipitation was the same as in seawater. Cl⁻ in excess of the marine estimate was assumed to have originated from anthropogenic sources, such as road salt, and was used to correct Na⁺ values on an equivalent basis. Values adjusted by these procedures are designated by asterisks.

ANC was defined from charge-balance considerations as $\sum \text{strong-base cations (C}_B\text{)} - \sum \text{strong-acid anions (C}_A\text{)}$:

$$\text{ANC} = [\text{Ca}^{2+} + \text{Mg}^{2+} + \text{Na}^+ + \text{K}^+] - [\text{SO}_4^{2-} + \text{NO}_3^- + \text{Cl}^- + \text{A}_s^-] \quad (1)$$

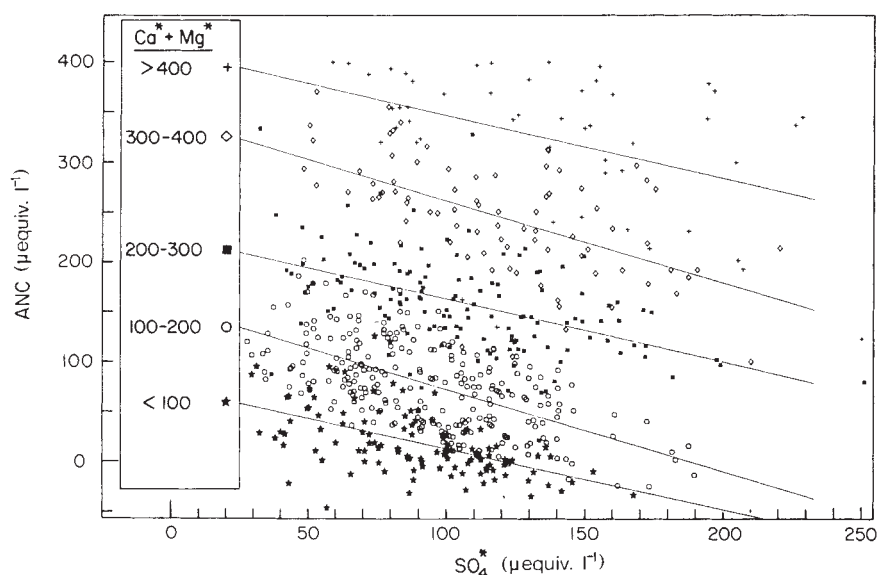


Fig. 2 Measured ANC plotted against non-marine sulphate for northeastern lakes having $\text{ANC} \leq 400 \mu\text{equiv. l}^{-1}$. Lakes are stratified according to $[\text{Ca}^* + \text{Mg}^*]$, and 15 lakes having $\text{SO}_4^* > 250 \mu\text{equiv. l}^{-1}$ were deleted. Best-fit linear regression lines are provided to indicate relative slope, but regression equations are not provided because they should not be used for predictive purposes.

where A_s^- refers to strong organic acid anion (all concentrations in $\mu\text{equiv. l}^{-1}$). We define 'acidic' lakes as those with measured (modified Gran analysis) $\text{ANC} \leq 0$.

When expressed as median values for each subregion, lakewater $[\text{SO}_4^{2-}]$ was highly correlated with estimated wet SO_4^{2-} deposition for each subregion (Fig. 1). Several processes can modify surface water $[\text{SO}_4^{2-}]$ relative to input. In addition, dry deposition of sulphur, which was not estimated because regional data are not yet available, can significantly alter the total sulphur input relative to measured wet deposition. It is therefore not surprising that scatter exists in a plot of lake $[\text{SO}_4^{2-}]$ against wet deposition. In spite of these uncertainties, however, lake sulphate concentrations in the western, upper midwestern and northeastern United States and Florida exhibit an approximately linear relation with increasing wet deposition. The Southern Blue Ridge subregion appears to be an outlier, supporting the observation that net sulphur retention (principally sulphate adsorption in soils) is high in many southeastern systems^{7,8}. The high variability of lake $[\text{SO}_4^{2-}]$ in Florida reflects the heterogeneity of lakes in this subregion. Extremely high sulphate concentrations in many Florida lakes appear to be associated with an influx of sulphur-rich groundwater from the Floridan aquifer³.

We did not find an obvious relationship between lake $[\text{SO}_4^*]$ and ANC for the full regional or subregional data sets, in agreement with analyses of data from eastern Canada⁹. This is because $[\text{SO}_4^*]$ is expected to correlate with loss of ANC, rather

than ANC *per se*. $[\text{Ca}^* + \text{Mg}^*]$ provides a better estimate of original lake susceptibility to acidic deposition effects than does ANC for low-base-cation lakes¹⁰. We therefore stratified northeastern lakes into five $[\text{Ca}^* + \text{Mg}^*]$ strata. A statistically significant ($P < 0.001$) negative relationship was found between lake $[\text{SO}_4^*]$ and ANC within each of the four lowest $[\text{Ca}^* + \text{Mg}^*]$ strata (Fig. 2). Multiple regression analysis of measured ANC as a function of $[\text{Ca}^*]$, $[\text{Mg}^*]$, and $[\text{SO}_4^*]$ yielded the following equation for lakes in the north-east:

$$\text{ANC} = 23.86 + 1.02[\text{Ca}^*] + 0.94[\text{Mg}^*] - 0.96[\text{SO}_4^*] \quad (R^2 = 0.98, n = 759) \quad (2)$$

Subregionally, results yielded comparable parameter estimates, with R^2 ranging from 0.91 (southern New England) to 0.99 (Adirondacks and central New England). The observed good fit reconfirmed our charge-balance definition of ANC presented in equation (1) (ignoring Na^+ , K^+ , NO_3^- , Cl^- , A_s^- and marine inputs of Ca^{2+} , Mg^{2+} and SO_4^{2-}). These results are illustrative of the negative relationship between $[\text{SO}_4^*]$ and ANC in northeastern lakes, when $[\text{Ca}^*]$ and $[\text{Mg}^*]$ are also taken into account.

A small number of lakes had very high $[\text{SO}_4^*]$, probably derived from watershed sources. This theory was tested by examining the ratio of $[\text{SO}_4^{2-}]$ in lakewater to $[\text{SO}_4^{2-}]$ in precipitation, using volume-weighted mean precipitation chemistry values interpolated to each lake site from NADP/NTN data for the 1-year period before lake sampling. Removal of SO_4^{2-} is thought to be low for northeastern lakes from SO_4^{2-} adsorption

Table 1 Lake chemistry from North American areas currently receiving low wet SO_4^{2-} deposition (generally $< 10 \text{ kg ha}^{-1} \text{ yr}^{-1}$)

Subregion*	n†	pH	Median			Percentage of lakes $\leq \text{ANC}$ criterion		
			ANC	SO_4^{2-} ($\mu\text{equiv. l}^{-1}$)	$\text{A}_s^- \ddagger$	0	20 ($\mu\text{equiv. l}^{-1}$)	40
NW Ontario	1,078	7.1	254	75	94	0.002	<2	4
Labrador	182	6.4	46	24	47	0	15	43
NE Minnesota	150	6.9	185	62	89	0	0	2
California	150	6.9	63	7	10	0	8	27
Pacific NW	159	7.0	140	15	13	0	6	17
Northern Rocky Mts	143	7.4	195	16	12	0	<1	8
Central Rocky Mts	129	7.2	105	24	12	0.6§	1	4
Southern Rocky Mts	139	7.6	317	35	15	0	<1	3

* Canadian data from Jeffries *et al.*⁹ and Jeffries (personal communication), Minnesota data from Linthurst *et al.*³, Western subregions from Landers *et al.*⁴. All summary statistics presented above, except data from Canada, are population estimates.

† n = number of lakes with measured ANC.

‡ A_s^- (organic anion concentration) estimated using method described by Oliver *et al.*¹⁷.

§ One naturally acidic lake sampled in Yellowstone National Park, influenced by a hot spring.

in watershed soils^{7,8} and microbial SO_4^{2-} reduction¹¹ in these primarily drainage systems. Thus, if dry deposition may be slightly in excess of wet deposition¹² and the evapoconcentration factor is near two¹³, we expect lake SO_4^{2-} :precipitation SO_4^{2-} ratios to have an upper bound of $\sim 4:1$ to $5:1$, although a definitive bound cannot be established. For northeastern lakes included in this study having $[\text{Ca}^* + \text{Mg}^*] \leq 200 \mu\text{equiv. l}^{-1}$, this ratio was generally within the expected range ($\text{mean} \pm 2 \text{ s.d.} = 2.92 \pm 1.60$). Many lakes with high base cation concentrations ($[\text{Ca}^* + \text{Mg}^*] > 200 \mu\text{equiv. l}^{-1}$) exhibited lake SO_4^{2-} enrichment considerably higher than expected (4.29 ± 4.98). For lakes with low base cation concentrations ($[\text{Ca}^* + \text{Mg}^*] \leq 200 \mu\text{equiv. l}^{-1}$), 2.1% of the population had a ratio exceeding 5.0, and 0.2% exceeding 6.0. For lakes with higher base cation concentrations, however, many more lakes exceeded these criteria, 23.1% and 13.7%, respectively. It appears that internal sources of SO_4^{2-} are significant for many of the lakes with high base cation and ANC concentrations. Although internal sources of SO_4^{2-} cannot be ruled out for lakes with low base cation concentrations, we observed no evidence that such sources are substantial.

Table 1 gives recent lake chemistry distributions from areas of North America currently receiving low wet deposition of SO_4^{2-} (generally $< 10 \text{ kg ha}^{-1} \text{ yr}^{-1}$). Despite widely varying ANC (median 46 to 317 $\mu\text{equiv. l}^{-1}$), estimated organic anion concentrations (median 10 to 94 $\mu\text{equiv. l}^{-1}$), geology, soils and vegetation, all subregions were remarkably similar in that they had few or no acidic lakes. Although several acidic lakes with high dissolved organic carbon were sampled in the Okefenokee Swamp of Georgia³, it appears that, on a regional basis, in temperate glaciated terrain, acidic lakes are generally absent in areas with low SO_4^{2-} deposition. We emphasize, however, that small lakes (especially those $< 4 \text{ ha}$ and to a lesser extent, those from 4 to 10 ha) are under-represented in the data from Canada and Minnesota (Table 1). Note also that the two pristine subregions containing the highest estimated organic anion concentrations (Northwest Ontario and Northeast Minnesota, 94 and 89 $\mu\text{equiv. l}^{-1}$, respectively) also had low percentages of low-ANC lakes. Only 4% and 2%, respectively, of the lakes in these subregions had $\text{ANC} \leq 40 \mu\text{equiv. l}^{-1}$.

Recognizing the inherent limitations of interregional comparisons because of differences in physical characteristics and deposition history, the data in Table 1 suggest that acidic northeastern lakes $> 4 \text{ ha}$ would likely not be acidic in the absence of SO_4^{2-} . This contention can be further examined on the basis of the $\text{SO}_4^{2-}/\text{C}_B^*$ ion ratio in acidic NLS lakes. From the charge balance definition of ANC (equation (1)), it follows that if SO_4^{2-} exceeds C_B^* , then lakewater will be acidic. Of the northeastern lakes measured as acidic (population estimate, $N = 326$), 85% also had a ratio of $\text{SO}_4^{2-}/\text{C}_B^* > 1$. This suggests that these lakes are now acidic due to SO_4^{2-} .

It is evident that most northeastern lakes that now have $\text{ANC} \leq 0 \mu\text{equiv. l}^{-1}$ are acidic because of SO_4^{2-} . It is not clear, however, whether some of these lakes may previously have been acidic from another source, such as organic acids^{14,15}. But available data from low-deposition areas in North America (Table 1) suggest that, for the lakes considered (that is, those generally $> 4 \text{ ha}$), temperate North American lakes are not acidic in the absence of anthropogenic sulphate. Palaeolimnological investigations of diatom remains in lake sediments have suggested that many northeastern lakes were acidic in pre-industrial times, but these have almost exclusively been small lakes ($< 10 \text{ ha}$)¹⁶.

We thank B. Rochelle, Roze Royce and R. Sauter for assistance; S. Christie, A. Henriksen, D. Johnson, E. Krug, D. McKenzie, H. Seip and P. Shaffer for valuable criticisms; A. Olsen and R. Olson for interpolating wet deposition chemistry to NLS sites; and D. Brown for lake disturbance classification. The research described here was funded as part of the National Acid Precipitation Assessment Program by the EPA. This paper was prepared at the EPA's Research Laboratory in Corvallis, Oregon, in part through Contract 68-03-3246 with Northrop,

Services, and is a contribution from EPA's Aquatic Effects Research Program. It has been subjected to the agency's peer and administrative review and approved for publication.

Received 27 August; accepted 17 November 1987.

- Drablos, D. & Tollan, A. (eds) *Proc. Int. Conf. Ecol. Impact Acid Precip.* (S.N.S.F. Project, Norway, 1980).
- Altshuler, A. P. & Linthurst, R. A. (eds) *The Acidic Deposition Phenomenon and its Effects: Critical Assessment Review Papers* Vol 2 (EPA, Washington, DC, 1984).
- Linthurst, R. A. et al. *Characteristics of Lakes in the Eastern United States* Vol. 1 (EPA, Washington, DC, 1986).
- Landers, D. H. et al. *Characteristics of Lakes in the Western United States* Vol 1 (EPA, Washington, DC, 1987).
- Watson, C. R. & Olsen, A. R. *Acid Deposition System (ADS) for Statistical Reporting: System Design and User's Code Manual* (EPA, Washington, DC, 1984).
- Hidy, G. M. & Young, J. R. *Acid Deposition and the West: A Scientific Assessment* (WEST Associates, Newbury Park, CA, 1986).
- Galloway, J. N., Norton, S. A. & Church, M. R. *Environ. Sci. Technol.* **17**, 541A (1983).
- Rochelle, B. P. & Church, M. R. *Water Air Soil Pollut.* **36**, 61–73 (1987).
- Jeffries, D. S., Wades, D. L., Kelso, J. R. M. & Linthurst, R. A. *Water Air Soil Pollut.* **31**, 551–567 (1986).
- Wright, R. F. *Environ. Sci. Technol.* (in the press).
- Rudd, J. W. M. et al. *Limnol. Oceanogr.* **31**, 1267–1280 (1986).
- Summers, P. W., Bowersok, V. C. & Stensland, G. J. *Water Air Soil Pollut.* **31**, 523–535 (1986).
- Geraghty, J. J., Miller, D. W., van der Leeden, F. & Troise, F. L. *Water Atlas of the United States* (Water Information Center, Port Washington, NY, 1973).
- Davis, R. B., Anderson, D. S. & Berge, F. *Nature* **316**, 436–438 (1985).
- Krug, E. C., Isaacson, P. J. & Frink, C. R. *J. Air Pollut. Control Ass.* **35**, 109–114 (1985).
- National Research Council *Acid Deposition Long-Term Trends* (National Academy Press, Washington DC, 1986).
- Oliver, B. G., Malcolm, R. L. & Thurman, E. M. *Geochim. cosmochim. Acta* **47**, 2031–2035 (1983).

CO₂ record in the Byrd ice core 50,000–5,000 years BP

A. Neftel, H. Oeschger, T. Staffelbach & B. Stauffer

Physikalisches Institut, Sidlerstrasse 5, CH-3012 Bern, Switzerland

The analysis of air in polar ice cores revealed 30% lower CO₂ values during glacial periods than during interglacial periods. At present this is confirmed by results from six different ice cores, two from Greenland and four from Antarctica^{1–5}. In all cores the CO₂ change coincides with the change in the isotopic composition of the ice, expressed as either the $\delta^{18}\text{O}$ or δD ratio; both are indicators for the mean annual surface temperature⁶. To investigate the relationship between atmospheric CO₂ concentration changes with changes in climate, the atmospheric CO₂ concentration during and at the end of the last glaciation has to be known in detail. To achieve this, we have studied a great number of samples from the deep ice core from Byrd station, Westantarctica, drilled in 1968. These measurements allow us to reconstruct the atmospheric CO₂ concentration in the time period 50,000–15,000 yr BP in great detail.

The samples from Siple Station, Westantarctica, allow us to study in detail the atmospheric CO₂ concentration of the past 200 years⁷. The ice record overlaps the Mauna Loa record for a short period, and so there are direct atmospheric measurements available for comparison. This is strong evidence that air stored in bubbles of polar ice samples reflect the true atmospheric composition. There are however several 'pitfalls' which can lead to deviations from the true atmospheric composition. Higher concentrations can be found, if melting and refreezing processes occur at the surface during the summer season². For the Byrd core no melt features were observed⁸. Carbon dioxide in the bubbles might also interact with impurities in the ice, particularly alkaline dust. We consider this effect to be of minor influence for the Byrd core, because the Byrd ice is slightly acidic along the whole core. The fact that the same CO₂-level changes at the glacial/interglacial transition are found in six different cores with different alkalinity, means that alkaline dust is not responsible for the drop in the CO₂-level. Regarding the different CO₂ irregularities measured in the depth interval 1,850–1,900 m below the surface of the Greenland Dye 3 core⁹ we cannot rule out the fact that CO₂ originating from dissolved carbonate dust

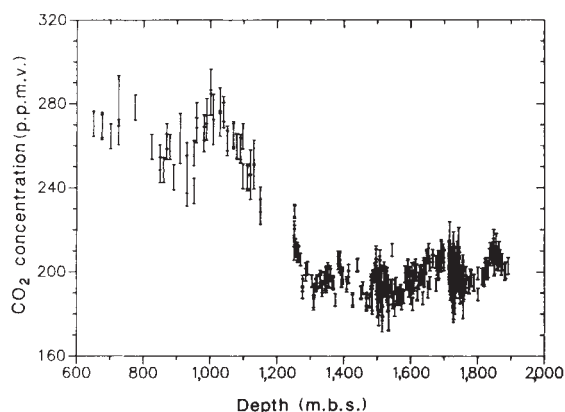


Fig. 1 All measured data points as a function of the depth below surface. The data points are indicated with a bar corresponding to the 1σ error.

may be partly or fully responsible for the higher CO_2 values. Other deviations might result from coring and post-coring effects. Samples which originate from brittle core sections show increased CO_2 concentrations². Such samples often have internal cracks filled with an organic fluid used to counterbalance the hydrostatic pressure in the drill hole. If the ice samples started to melt during storage or transport, followed by a refreezing, the CO_2 content in the bubbles could be depleted in some parts of the ice core and enriched in others.

Because the enclosure process acts as a low-pass filter, the CO_2 record stored in the bubbles of the polar ice archive is a smoothed record of the atmospheric CO_2 concentration. In the Byrd core the air is enclosed between 60 and 80 m below the surface (m.b.s.) and the duration of the enclosure is ~ 50 yr during the Holocene¹⁰. The analysis of small adjacent samples allows us to discriminate to a certain extent, variations of CO_2 in the ice caused by natural atmospheric variations from 'artefacts' (as described above). For non-natural changes, local CO_2 concentration gradients are expected to be larger and to vary irregularly when compared to natural atmospheric variations, which are smoothed by the enclosure process. Our system has the advantage that we can measure small samples (2–10 g) and therefore we can detect local variations indicating non-atmospheric disturbances.

We report here a new series of measurements of the Byrd core consisting of samples from the depth interval 600–1,900 m below the surface. In the depth interval 600–1,000 m.b.s. the core quality is rather poor. The ability to measure very small samples means that we can avoid samples with visible internal cracks. Unfortunately no more core material was available in the depth interval 1,150–1,250 m.b.s. Figure 1 shows all the measured data as a function of the depth. The results agree generally with earlier published data². Different timescales for the Byrd core are reported in the literature^{11–13}. Our interpretation is based on the most recent dating by Hammer *et al.*¹³. It relies on a continuous measurement of the acidity along the whole core. The depth interval 600–1,900 m.b.s. corresponds approximately to the time interval 6,000–50,000 yr BP. The 'missing' depth interval 1,150–1,250 m.b.s. corresponds to an important part of the transition from the last glaciation to the Holocene. During this period fast climatic changes have been reported from different records from the Northern Hemisphere^{14,15}. To calculate the atmospheric concentration by deconvolution, with the exception of the depth interval 1,150–1,250 m.b.s. we interpolated the data to obtain an equidistant time series. The mean sampling rate of the measurements corresponds to one sample every 200 years. We used the following expression to describe the enclosure process:

$$c_{\text{ice}}(t) = \int_{-\infty}^{\infty} f(t' - t) c_{\text{atm}}(t') dt' \quad (1)$$

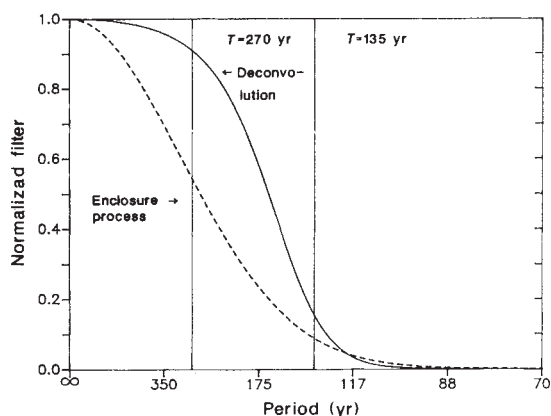


Fig. 2 Filter characteristic of the enclosure process (dotted line) and the deconvolution algorithm (solid line). The vertical lines indicate the periods corresponding to the close-off time and twice the close-off time respectively for ice age conditions.

where $f(t' - t)$ is a normalized gaussian filter:

$$f(t' - t) = \frac{1}{(\pi\tau^2)^{1/2}} \exp\left[-\frac{(t' - t)^2}{\tau^2}\right] \quad (2)$$

Eighty-four per cent of the gas is enclosed in the time interval 2τ . For the Holocene (10,000–0 yr BP) we assumed τ to be 27 years¹⁰. There is strong evidence that the yearly accumulation decreased between 15,000 and 10,000 yr BP by a factor of 2.5. We assumed therefore that τ increases linearly for this period by the factor of 2.5 up to 67 years^{10,16}. The atmospheric concentration has to be calculated from equation (1). The deconvolution is calculated in the frequency domain. Convolution in the time domain corresponds to a multiplication in the frequency domain. The atmospheric concentration is then given by:

$$c_{\text{atm}}(t) = F^{-1}(c_{\text{atm}}(w)) = F^{-1}\left(\frac{c_{\text{ice}}(w)}{f(w)}\right) \quad (3)$$

F^{-1} is the inverse Fourier transform and $f(w)$ is the Fourier transform of the enclosure function. To achieve meaningful solutions the inverse filter $f^{-1}(w)$ was slightly modified to suppress the blow-up of high frequencies.

$$f_{\text{mod}}^{-1}(w) = \frac{\exp((-w\tau/2)^2)}{\exp(-2(w\tau/2)^2) + \exp(-2(g/2)^2)} \quad (4)$$

In equation (4) g was set to a value of 2.5. The filter characteristic of the deconvolution procedure is equal to the ratio of the modified enclosure function and the original enclosure function in the frequency domain. Figure 2 shows the filter characteristic of the deconvolution procedure. The solid line represents the overall characteristic of the deconvolution, the dotted line the characteristic of the enclosure process of the ice. Oscillations of the atmospheric CO_2 concentration with a period corresponding to twice the enclosure time, 2τ would be attenuated to 40% in the ice and would be reinstated to 82% of the original value after the deconvolution procedure. For oscillations with a period corresponding to the duration of the enclosure time, the percentages would be 8.5% for the CO_2 record in ice and 18% for the reconstructed record by the deconvolution procedure. Faster changes are suppressed and cannot be seen either in the ice or reconstructed by deconvolution. The limited accuracy of the measured data points causes a certain uncertainty in the reconstructed atmospheric CO_2 concentration. We determined a band empirically by superposing 50 different records. For each record the deconvolution was calculated allowing the single measurements to take a random value within the estimated error limit. As a weight function a gaussian distribution was chosen. The band was then calculated as the 2σ interval of the mean curve

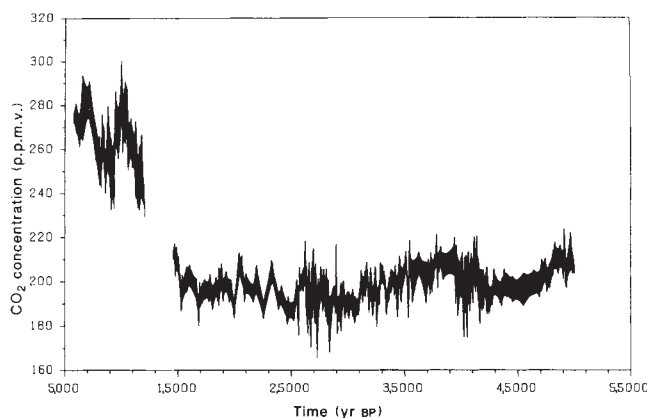


Fig. 3 Calculated band of the atmospheric CO₂ concentration.

where σ is the standard deviation. This band is plotted in Fig. 3.

We discuss the Byrd CO₂ record in two parts. The first part 15,000–5,000 yr BP includes the transition from glacial to post-glacial time. By bridging the missing part with a linear interpolation the CO₂ increase from 200 p.p.m.v. in the glacial period to 280 p.p.m.v. in the Holocene occurred in ~5,000 yr. The value of 280 p.p.m.v. which is reached at 10,000 yr BP corresponds to the pre-industrial value. We consider the observed minimum of 245–255 p.p.m.v. ~8,000 yr BP to be real and not due to artefacts related to the poor core quality. Our data indicate that at least at the beginning of the Holocene changes up to 40 p.p.m.v. of the atmospheric CO₂ concentration occurred. At present we have no explanation for such atmospheric CO₂ variation at the beginning of the Holocene. Figure 4 shows a comparison of the CO₂ record with the $\delta^{18}\text{O}$ record¹⁷ in the depth interval 1,000–1,600 m.b.s. covering the beginning of the transition from the last glacial age into the Holocene. The $\delta^{18}\text{O}$ record is shifted by 12 m towards greater depth to compare the same age for both records. This assumes an age difference of 600 years between the ice and the mean age in the bubbles. The $\delta^{18}\text{O}$ ratio starts to increase between 200 and 1,200 years before the CO₂ concentration starts to increase. At the time the CO₂ starts to rise, the $\delta^{18}\text{O}$ ratio increase is already 20–30% of the total shift at the glacial/interglacial transition. According to these results the atmospheric CO₂ therefore cannot have triggered the climatic shift observed in the Byrd core, but it could be an important forcing factor establishing the mild climate of the present interglacial.

The second part 50,000–15,000 yr BP shows variations in a band between 180 and 220 p.p.m.v. These variations found in the Byrd core are much smaller than those observed in the Dye 3 core, which are characterized by 50 p.p.m.-elevated CO₂ concentrations for several relatively short periods between 40,000 and 30,000 yr BP.

CO₂ is globally well mixed. The interhemispheric gradient cannot exceed a few p.p.m. Therefore for corresponding gas ages the same concentration must be measured in cores from both hemispheres, if they represent the atmospheric concentration. It is straightforward to check whether a given pattern believed to be atmospheric can be found in different cores. If not, the specific pattern cannot be atmospheric.

There is a contradiction between the CO₂ record from the Byrd deep ice core and from the Dye 3 deep ice core. We see three possible explanations: (1) The elevated CO₂ concentrations in the Dye 3 core are caused by melt features. (2) The duration of the warm periods with higher CO₂ levels is shorter than the gas enclosure process at Byrd station. (3) The dating of at least one core is completely wrong, so that entirely different time sequences are compared in both cores.

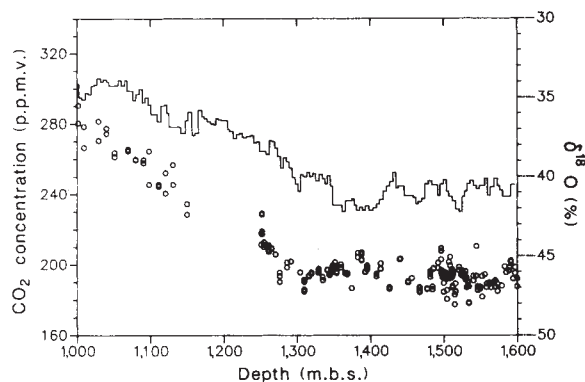


Fig. 4 $\delta^{18}\text{O}$ and CO₂ record 1,000–1,600 m.b.s. The solid line corresponds to the age-adjusted $\delta^{18}\text{O}$ curve.

The comparison with the Vostok record recently published by Barnola *et al.*⁴ shows a good agreement for the long-term variation and support the used dating of the two Antarctic cores.

Two possible mechanisms for the change in atmospheric CO₂ concentration which occurred during the transition from glacial time to interglacial time have been suggested. Broecker¹⁹ first pointed to a redistribution of carbon species in the ocean due to deposition of nutrient-rich sediments on continental shelves during the deglaciation period. With such a mechanism only slow changes as suggested, for example, by a smooth interpolation of the CO₂ data from the Byrd core can be explained. Faster changes, as suggested by CO₂ data from the ice age part of the Dye 3 core⁹ could be explained by another hypothesis including changes of circulation or biological productivity in the high-latitude ocean^{19,20}.

Variations during the glaciation period, similar to those found in polar ice core data have been reported from sea-sediment records²¹, indicating changes in the ocean biota system. Because of difficulties in dating both the ice sea-sediment cores we cannot yet correlate these variations quantitatively. More polar ice cores from cold locations with good core quality are needed to establish the exact phase relation between $\delta^{18}\text{O}$ and CO₂ during the glacial/interglacial transition and to characterize quantitatively the small CO₂ variations in the glaciation period.

We thank Dr C. C. Langway for providing us with the Byrd ice core samples. This work was supported by the Swiss NSF, the Division of Polar Programs of the US NSF and the US Department of Energy.

Received 7 May; accepted 25 November 1987.

1. Berner, W., Oeschger, H. & Stauffer, B. *Radiocarbon* **22**, 227–235 (1980).
2. Neftel, A., Oeschger, H., Schwander, J., Stauffer, B. & Zumbund, R. *Nature* **295**, 220–223 (1982).
3. Delmas, R. J., Ascencio, J. M. & Legrand, M. *Nature* **284**, 155–157 (1980).
4. Barnola, J. M., Raynaud, D., Korotkevich, Y. S. & Lorius, C. *Nature* **329**, 408–413 (1987).
5. Stauffer, B., Neftel, A., Oeschger, H. & Schwander, J. *Geophys. Monogr.* **32**, 85–94 (1985).
6. Dansgaard, W. *Tellus* **16**, 430–468 (1964).
7. Neftel, A., Moor, E., Oeschger, H. & Stauffer, B. *Nature* **315**, 45–47 (1985).
8. Gow, A. J. *CRREL Res. Rep.* **197**, 7 (1968).
9. Stauffer, B., Hofer, H., Oeschger, H., Schwander, J. & Siegenthaler, U. *Ann. Glaciol.* **5**, 160–164 (1984).
10. Schwander, J. & Stauffer, B. *Nature* **311**, 45–47 (1985).
11. Johnson, S. J., Dansgaard, W., Clausen, H. B. & Langway, C. C. Jr *Nature* **235**, 429–434 (1972).
12. Thompson, L. G., Hamilton, W. L. & Bull, C. J. *Glaciol.* **14**, 433–444 (1975).
13. Hammer, C. U., Clausen, H. B. & Langway, C. C. Jr *Ann. Glaciol.* **7**, 214 (1985).
14. Dansgaard, W. *et al. Geophys. Monogr.* **29**, 288–298 (1984).
15. Oeschger, H. *et al. Geophys. Monogr.* **29**, 299–313 (1984).
16. Herron, M. M. & Langway, C. C. Jr *J. Glaciol.* **25**, 373–385 (1980).
17. Robin, G. de Q. *The Climatic Record in Polar Ice Sheets* 113–118 (Cambridge University Press, 1983).
18. Broecker, W. S. *Progr. Oceanogr.* **11**, 151–197 (1982).
19. Siegenthaler, U. & Wenk, T. *Nature* **308**, 624–625 (1984).
20. Sarmiento, J. L. & Toggweiler, J. R. *Nature* **308**, 621–624 (1984).
21. Ruddiman, W. F. & McIntyre, A. *Paleogeogr., Paleoclimatol. Paleocool.* **35**, 145–214 (1981).

Impact frustration of the origin of life

Kevin A. Maher & David J. Stevenson

Division of Geological and Planetary Sciences,
California Institute of Technology, Pasadena,
California 91125, USA

One possible definition for the origin of life on Earth is the time at which the interval between devastating environmental insults by impact exceeded the timescale for establishing self-replicating proto-organisms. A quantitative relationship for the Hadean (pre-3,800 Myr ago) and Early Archean (3,800 to 3,400 Myr) impact flux can be derived from the lunar and terrestrial impact records. Also, the effects of impact-related processes on the various environments proposed for abiogenesis (the development of life through chemical evolution from inorganic materials) can be estimated. Using a range of plausible values for the timescale for abiogenesis, the interval in time when life might first have bootstrapped itself into existence can be found for each environment. We find that if the deep marine hydrothermal setting provided a suitable site, abiogenesis could have happened as early as 4,000 to 4,200 Myr ago, whereas at the surface of the Earth abiogenesis could have occurred between 3,700 and 4,000 Myr.

As a result of the lack of a Hadean terrestrial rock record, many of the constraints on the mode, environment and timing of abiogenesis are derived from laboratory simulations or from theoretical extrapolations to early terrestrial conditions, either forward in time from the formation of the Solar System or back from the Archean record¹. In recent years the modification of biological evolution by impacts has received increasing attention, primarily because of the large impact that apparently occurred at the Cretaceous/Tertiary boundary^{2,3}. The role of impacts at earliest times can be roughly quantified because of our knowledge of the lunar impact chronology^{4,5} during the Late Heavy Bombardment and of the terrestrial cratering record over the last 2×10^9 yr⁶⁻⁸. Because of the inevitable uncertainties, we present our modelling more as an indicator of the level of environmental insult to the early Earth, rather than as a basis for precise prediction.

Most current theories for the origin of life have a common theme of chemical evolution, starting with the ordinary constituents of the atmosphere, crust and hydrosphere in some reducing environment, and resulting in chemoautotrophic bacteria-like organisms. The Earth at this time apparently possessed a thick neutral ($\text{CO}_2 + \text{N}_2$) secondary atmosphere⁹ and extensive oceans, and was probably quite warm¹⁰ from a CO_2 -greenhouse effect, despite decreased solar luminosity¹¹, and had little or no cratonic crust¹², though some form of crustal differentiation¹³ and plate tectonics¹⁴ was probably taking place. Bacterial fossils in the rock record show that by 3,500 Myr the biosphere already existed¹⁵.

Within this framework, several environments have been proposed as candidates to host abiogenesis. One is an evaporitic lake or lagoon or shallow marine area¹⁶ (the 'warm puddle') where there would have been a dilute soup of organics and energy sources for low-weight organic molecule formation (lightning, ultraviolet radiation, evaporation). A second possibility is the hydrothermal spring systems of the mid-ocean ridges or their Hadean analogues, which would have had high thermal and chemical gradients, an abundance of catalysts and substrates for the growth of more complex organic molecules, and convective flow to move products through the system¹⁷. A third potential site would be subaerial: the soil or shallow weathered subsurface, where the first self-replicating 'organisms' could have been clay minerals capable of mutation, natural selection and growth, and of acting as a substrate to the first self-replicating organic molecules¹⁸. The timescale (T_a) for abiogenesis under these

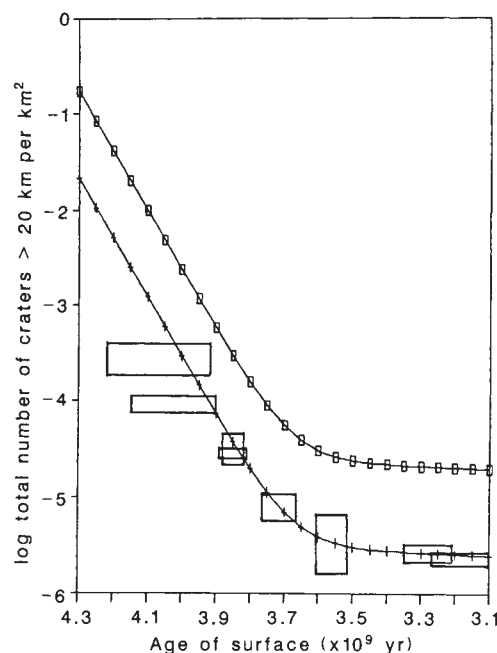


Fig. 1 The lunar and terrestrial cratering history for total number of craters ≥ 20 km per km^2 as a function of surface age, determined from the relationships in the text. Saturation is neglected. The open boxes represent data from crater surveys from the Moon⁴. The lunar curve was calculated from the terrestrial, assuming a correlation factor of 8 (a result of gravitational focusing and different impact velocities and surface gravities). Symbols: squares, calculated terrestrial; crosses, calculated lunar.

different circumstances is not known, but is potentially quite short¹⁹. We have chosen plausible values of 10^5 to 10^6 years for the hydrothermal model, 3×10^5 to 3×10^6 for the warm puddle case, and 10^6 to 10^7 for the soil hypothesis. We acknowledge that these are highly uncertain, but even large changes in these values will not substantially alter the results.

From the lunar cratering record^{4,5} we adopt a simple model for the impact flux $f(t)$:

$$f(t) = 1 + \exp[(t - t_0)/T]$$

where t is the time measured backward from the present. We assume $t_0 = 3,400$ Myr and the decay constant T is 70 Myr. For $N(D)$, the cumulative number of craters $\geq$ diameter D per unit area per unit time, we adopt^{20,21}

$$N(D) = k \cdot (1 \text{ km}/D)^{1.8} \cdot f(t) \text{ km}^{-2} \text{ yr}^{-1}$$

where k is a constant determined from the terrestrial cratering history of the last several billion years⁶⁻⁸ as $\sim 1.4 \times 10^{-12}$. This is also consistent with the lunar record, after allowance for gravitational focusing, different impact velocities and surface gravities. The corresponding incremental number density $n(D)$ is defined by

$$n(D) = -dN/dD = 1.8 \cdot k \cdot (1 \text{ km}/D)^{2.8} \cdot f(t) \text{ km}^{-3} \text{ yr}^{-1}$$

The total number of cratering events $M(D)$ (neglecting saturation) by the age of a surface (see Fig. 1) is

$$M(D) = \int_0^t k \cdot f(t) \cdot (1 \text{ km}/D)^{1.8} dt$$

$$M(D) \sim k \cdot (1 \text{ km}/D)^{1.8} [t + T \cdot \exp\{(t - t_0)/T\}]$$

We assume that d_{max} , the maximum impactor diameter, was of the order of 1,000 km and would have caused a crater (D_{max}) of about 2,600 km diameter, consistent with both the known population of the asteroid belt (which may represent a frozen sample of the impacting spectrum²²) and with the lunar basins. We relate the crater diameter (D) and impactor diameter (d)

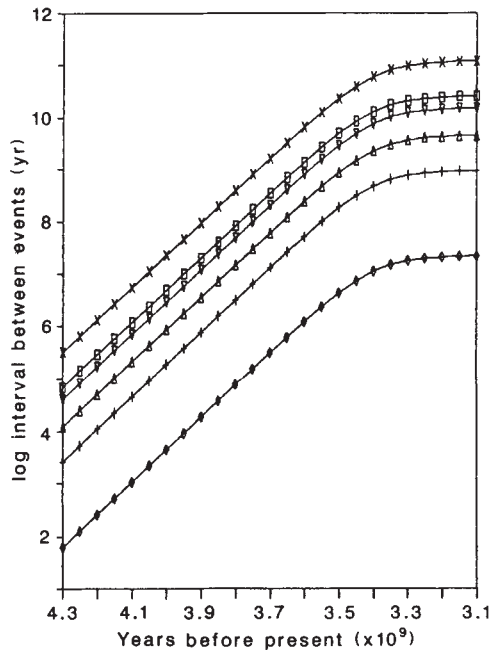


Fig. 2 The interval between impact-related environmental trauma as a function of time. Symbols: squares, cratering at the surface; upside-down triangles, cratering on ocean bottom; daggers, surface cratering with ejecta thickness of 3 m; diamonds, global surface climatic trauma; triangles, global surface sterilization; crosses, possible sterilization of whole planet.

(ref. 23)

$$D = 3.31 \cdot V^{1/3} \cdot d^{5/6} \text{ km}$$

with the caveat that the applicability to large impacts is questionable. V is the impact velocity (here taken to be 13–15 km sec⁻¹). $D_{\min}$, the minimum crater size, would be about 120 m due to atmospheric resistance.

Land-based impact events excavate a bowl-shaped crater whose size is primarily dictated by kinetic energy. Later crater enlargement by collapse of 35% (ref. 24) is well within the region of severe effects and will not be further considered. The resulting ejecta has a thickness l (refs 25, 26)

$$l = 23.2 \cdot (R/1 \text{ km})^{4.24} \cdot (1 \text{ km}/r)^{3.5} \text{ m}$$

where r is the distance from the centre of a crater of radius R . The area of diameter D_e covered by ejecta to a thickness l or greater for a crater of diameter D is then

$$D_e = 2.12 \cdot (1 \text{ m}/l)^{0.286} \cdot (D/1 \text{ km})^{1.21} \text{ km}$$

The incremental size-frequency spectrum $n_e(D_e)$ for ejecta blankets is given by

$$n_e(D_e) = 4.56 \cdot k \cdot f(t) \cdot (1 \text{ m}/l)^{0.425} \cdot (1 \text{ km}/D_e)^{2.49} \text{ km}^{-3} \text{ yr}^{-1}$$

Besides direct impact excavation and shock, the ejecta blankets would seriously disrupt the chemical evolution processes in the 'warm puddles' as a result of heating, physical disruption and mixing, covering of the sediment-water interface, and acidification. Soil and shallow subsurface environments would also be negatively affected by burial, causing changes in chemical concentration and pH, groundwater flow and heating.

Pelagic impacts would behave substantially differently from land-based impacts, with greatly increased shock vaporization and injection of water and dust into the atmosphere, catastrophic tsunamis, a subdued or non-existent benthic crater topography, and a relative lack of impact ejecta^{27,28}. For projectiles less than about 1.5 km in size (a crater of 11 km) in an ocean of 3 km depth, significant benthic excavation does not seem to occur,

due to shielding by the hydrosphere. We adopt this as the threshold for significant trauma to the ocean floor, and assume that an area twice that of the crater will be negatively affected by scouring from the returning water column and shock waves.

Sufficiently large impacts would have been felt more widely. The giant tsunamis generated could have had truly global and catastrophic results²⁸. Shallow marine and land surfaces might have been wiped clean, whereas deeper marine environments would have been less disturbed. Immediately after the impact, shock heating would have driven up surface temperatures, but loss of sunlight due to material injected into the atmosphere might cause a drop to below freezing for the next year or so, with light levels far below those necessary for photosynthesis³⁰. The surface of the Earth might also have been extensively acidified due to the shock formation of nitrogen-bearing compounds³¹. We expect that the minimum projectile diameter for severe climatic trauma is similar in size to the body proposed to have caused the Cretaceous/Tertiary extinction event². The deep ocean would apparently be shielded against all but the greatest of such events (N. Sleep, personal communication). The possibility of impacts large enough to totally sterilize the surface, or even the ocean depths as well, must be considered. Bodies of the order of 65 km (crater of 265 km) could heat up the atmosphere and surface by 100 °C. Similarly, those 250 km in size (crater of 850 km) might sterilize the whole planet including the ocean bottoms (N. Sleep, personal communication). An exception could possibly be heat-tolerant thermophilic bacteria, which might preferentially survive such heating (S. Hoffman, personal communication), but not perhaps the accompanying ejecta fouling and water column mixing.

With these threshold values for impact-induced trauma, the time interval between such events (T_{imp}) can be found and compared to T_a . T_{imp} includes direct crater obliteration, T_c ; trauma by the associated ejecta debris blanket, T_e ; world-wide events due to impact, T_w ; and events on the ocean bottom, T_b . The timescales are found by the following relationships (R_E is the radius of the Earth):

$$T_c^{-1} = \int_{D_{\min}}^{D_{\max}} n(D) \cdot (\pi D^2/4) dD \text{ yr}^{-1}$$

$$T_e^{-1} = \int_{D_{e,\max}}^{D_{e,\min}} n_e(D_e) \cdot (\pi D_e/4) dD_e \text{ yr}^{-1}$$

$$T_w^{-1} = n(D) \cdot (4\pi R_E^2) \text{ yr}^{-1}$$

$$T_b^{-1} = \int_{D_{\max}}^{D_{\min}} 2 \cdot n(D) \cdot (\pi D^2/4) dD \text{ yr}^{-1}$$

This follows directly from the definitions of $n(D)$ and $n_e(D_e)$ and the realization that the right-hand side in the above equation is a probability per unit time that any point on the Earth's surface is within the area of actual impact or corollary events. Thus

$$T_c = 1.0 \times 10^{11} \cdot [f(t)^{-1}] \cdot [(D_{\max}^{0.2} - D_{\min}^{0.2})^{-1}] \text{ yr}$$

$$T_e = 1.0 \times 10^{11} \cdot [f(t)^{-1}] \cdot [(D_{e,\max}^{0.51} - D_{e,\min}^{0.51})^{-1}] \text{ yr}$$

$$T_w = 2.5 \times 10^2 \cdot [f(t)^{-1}] \cdot (D_{\min}/1 \text{ km})^{2.8} \text{ yr}$$

$$T_b = 5.0 \times 10^{10} \cdot [f(t)^{-1}] \cdot [(D_{\max}^{0.2} - D_{\min}^{0.2})^{-1}] \text{ yr}$$

These timescales are shown in Fig. 2. The results are not very sensitive to the values used for $D_{\max}$ and $D_{\min}$.

In accordance with our basic thesis, the time at which the impact time scale T_{imp} for an environment was equal to T_a defines the first point at which that environment would have gone long enough untraumatized that abiogenesis could have taken place. We assume that life would have appeared almost as soon as it could have. Although statistically there would have been some patches that would be left undisturbed long enough earlier, the exponential increase in impact flux would cause our estimates to be insensitive to these exceptions.

Our calculations give rise to some interesting implications for the origin and early development of life on Earth. Although there are minor differences among different surface environments, we conclude that, for an ejecta blanket of 3 m as the criterion for serious disruption, the first primitive organism might have evolved at the surface between 4,000 and 3,700 Myr. Estimates for different assumptions about T_e and critical ejecta thickness can be found to be in the same range. In the deep ocean hydrothermal vents, the origin of life could have taken place as early as 4,200 to 4,000 Myr—substantially earlier than at the surface. In addition, the surface would probably have been sterilized as late as 3,600 to 3,700 Myr by global trauma due to impact processes, and even the deepest ocean environments might have been made terminally uncomfortable for life (except perhaps thermophilic forms) as late as 4,000 to 3,900 Myr.

Another factor is that severe climatic events would remain frequent (every 10^5 to 10^6 yr) until 3,600 to 3,800 Myr. Of course, if the reader has a different preference for the timescale of abiogenesis, he can choose his own estimate for these events from Fig. 2.

Some conclusions can be drawn from these inferences. If life could have evolved in or near the mid-ocean ridge hydrothermal systems, then it probably began there. Because of the interconnected nature of the ridges versus the more patchy nature of the suitable surface sites, it might have been more likely to propagate as well. Wherever life did first appear, it would seem possible that it was eradicated from at least the surface of the planet (perhaps several times), re-evolving in some new location or radiating from a preserved, more heat-tolerant population each time, before it took possession of the Earth undisputed by impact events. These multiple extinctions (and possibly origins) are inferred from the extent of overlap between the period at which life first would have appeared and the last probable instance of impact sterilization for any of these environments. An additional conclusion is that photosynthesis would have been rather difficult before 3,700 to 3,800 Myr because of frequent climatic events, including a near-total lack of sunlight for extended periods. It is also intriguing that the biosphere contained a relatively complex and diverse set of organisms by 3,500 Myr in light of the potentially short time period involved since the last sterilization event, and/or inception of life. This could imply a type and rate of evolutionary process no longer in operation.

Received 11 August; accepted 3 December 1987.

- Chang, S. *Phys. Earth planet. Inter.* **29**, 261–280 (1982).
- Alvarez, L. W. *et al. Science* **208**, 1095–1108 (1980).
- Alvarez, W. *EOS* **65**, 649–658 (1986).
- Carr, M. H. *et al. NASA, SP-469* (1984).
- Shoemaker, E. M. in *The New Solar System* (ed Beatty, J. K. *et al.*) 33–44 (Cambridge Univ. Press, 1982).
- Shoemaker, E. M. *A. Rev. Earth planet. Sci.* **11**, 461–494 (1983).
- Grieve, R. A. F. *A. Rev. Earth planet. Sci.* **15**, 245–270 (1987).
- Grieve, R. A. F. & Dence, M. R. *Icarus* **38**, 230–242 (1979).
- Levine, J. S. & Augustsson, T. R. *Origins of Life* **15**, 299–318 (1985).
- Karhu, J. & Epstein, S. *Geochim. cosmochim. Acta* **50**, 1745–1756 (1986).
- Kasting, J. E. *Origins of Life* **14**, 75–82 (1984).
- Moorbath, S. *Origins of Life* **15**, 251–261 (1984).
- Froude, D. O. *et al. Nature* **304**, 616–618 (1983).
- Abbott, D. & Hoffman, S. *Tectonics* **3**, 429–448 (1984).
- Awramik, S. M. *et al. in Mineral Deposits and the Evolution of the Biosphere* (ed. Holland, H. D. *et al.*) 309–320 (Springer, New York, 1982).
- Lazcano, A. *et al. Precamb. Res.* **20**, 259–282 (1983).
- Baross, J. A. & Hoffman, S. E. *Origins of Life* **15**, 327–345 (1985).
- Cairn-Smith, A. G. *Scient. Am.* **256**, 90–100 (1985).
- Miller, S. L. in *Mineral Deposits and the Evolution of the Biosphere* (ed. Holland, W. D. *et al.*) 155–176 (Springer, New York, 1982).
- Shoemaker, E. M. & Morris, E. C. *Icarus* **12**, 188–212 (1970).
- Grieve, R. A. F. & Parmentier, E. M. *Proc. 27th Int. Geol. Cong.* **19**, 99–114 (1984).
- Hartman, W. K. in *Comets, Asteroids, Meteorites* (ed. Delsemme, A. H.) 277–281 (Univ. Toledo Press, 1977).
- Schmidt, R. M. & Holsapple, K. A. *Geol. Soc. Am. Sp. Pap.* **190**, 93–102 (1982).
- Shoemaker, E. M. *et al. in Asteroids* (ed. Gehrels, T.) 253–292 (Univ. Arizona Press, Tucson, 1979).
- McGetchin, T. R. *et al. Earth planet. Sci. Lett.* **20**, 226–236 (1973).
- Seebaugh, W. R. in *Impact and Explosion Cratering* (ed. Roddy, D. J.) 1043–1056 (Pergamon, New York, 1977).

- Melosh, H. J. *Geol. Soc. Am. Sp. Pap.* **190**, 121–127 (1982).
- Gault, D. E. & Sonett, C. P. *Geol. Soc. Am. Sp. Pap.* **190**, 69–92 (1982).
- O'Keefe, J. D. & Ahrens, T. J. *Nature* **298**, 123–127 (1982).
- Toon, O. B. *et al. Geol. Soc. Am. Sp. Pap.* **190**, 187–200 (1982).
- Prinn, R. G. *EOS* **46**, 813 (1985).

Thermoluminescence dating of Mousterian 'Proto-Cro-Magnon' remains from Israel and the origin of modern man

H. Valladas*, J. L. Reyss*, J. L. Joron†, G. Valladas*, O. Bar-Yosef‡ & B. Vandermeersch§

* Centre des Faibles Radioactivités, Laboratoire mixte CNRS-CEA, Avenue de la Terrasse, 91198 Gif sur Yvette Cedex, France

† Group des Sciences de la Terre, Laboratoire Pierre Süe, CEN, Saclay, 91191 Gif sur Yvette, France

‡ Isotope Department, Weizmann Institute of Sciences, Rehovot 76100, Israel

§ Laboratoire d'Anthropologie, Université de Bordeaux 1, UA 376 du CNRS, 33405 Talence, France

The Qafzeh¹ and Skhul^{2,3} caves in Israel have yielded the remains of over 30 hominids. Despite their association with Mousterian deposits, these have been recognized as forerunners of other *Homo sapiens sapiens* on morphological grounds and have been called 'Proto-Cro-Magnons'¹. Other west Asian caves (Amud, Tabun, Kebara and Shanidar) have yielded skeletal remains of Neanderthals associated with similar Middle Palaeolithic deposits. The lack of precise dates for these deposits has made it difficult to ascertain which of the two hominids was present first in the area^{1,4–9}. Recently we reported an age of ≈ 60 kyr for the Neanderthal burial at Kebara⁶ (Israel). Here we report thermoluminescence dates for 20 specimens of burnt flints recovered from the hominid-bearing layers of Qafzeh¹. The dates, which range from ~ 90 to 100 kyr BP, provide an independent measure for the great antiquity of southwest Asian modern humans which have previously been dated to ~ 40 kyr BP on the basis of European models^{7–9}. Our results also exclude a close phylogenetic relationship between the Cro-Magnons and Neanderthals^{1,4–6}.

The Qafzeh cave (lower Galilee) was excavated by Neuville and Stekelis (1932–1935) and more recently by Vandermeersch (1965–1979)^{1,10}. The 4.5 m thick Mousterian sequence is subdivided into two major accumulation units. The first (layers V–XV) is 2 m thick and is rich in broken mammalian bones and lithic artefacts, demonstrating intensive human occupation. The second (layers XVII–XXIV) is 2.5 m thick and is rich in bones of microvertebrates, indicating ephemeral human use of the cave. The rodent assemblages^{11–13} of the second accumulation unit, from which all the hominid remains have come, include two archaic species of African rats (*Mastomys batei* and *Arvicanthis ectos*) and a subspecies of the Eurasian dormouse (*Myomimus roachi qafzensis*), all of which became extinct during the early Mousterian. The presence of new arrivals, such as *Cricetulus migratorius* (a grey hamster), and the evolution of the modern dormouse subspecies (*Myomimus roachi roachi*) mark the late Mousterian deposits such as Tabun C, Hayonim cave upper E., Kebara, Geula¹³.

The cultural stratigraphy of the Tabun cave⁷ is often taken as a yardstick for Levantine Middle Palaeolithic sites. Artefact morphology and the thickness/width ratio of flakes indicate that the Qafzeh Mousterian resembles that of Tabun C¹⁴ or later assemblages and should therefore be dated to the end of the Middle Palaeolithic, that is ~ 40 kyr BP^{7,15}. The discrepancy between relative dates based on metrical lithic analysis (suggesting that the Neanderthals predated the Qafzeh hominids) and those derived from the seriation of microvertebrate assemblages

Table 1 Thermoluminescence results and radioactivity data for the flints from Qafzeh

Sample	U*	Th*	K*	Sα†	Dα‡	Dβ§	Internal dose	External dose	Annual dose		
	(p.p.m.)						(10 ⁻⁵ Gy yr ⁻¹)			Palaeodose (10 ⁻² Gy)	Age (kyr BP)
Layer XVII											
13	0.762	0.038	290	1.88	26.6±1.8	13.6	40.8±3	24.3	65±5	6,149±325	94.3±8.8
14	0.859	0.043	550	1.51	24.2±2.7	17.1	41.9±3.7	24.5	66±5	7,044±272	106.0±9.6
29	1.922	0.051	450	1.21	43.1±3.4	31.9	76.2±5.7	24.4	101±7	10,79±463	107.2±8.8
33	2.239	0.088	600	2.27	94.7±7.8	37.9	135±11	24	159±12	14,16±797	89.2±8.4
34	1.058	0.04	430	1.63	32.0±2.6	19.1	51.9±4	24.2	76±6	6,692±227	87.8±7.2
36	1.238	0.032	260	1.29	29.6±1.9	20.3	50.4±3.6	24.5	75±5	7,556±282	100.7±8.2
Layer XVIII											
38	1.314	0.086	960	2.13	52.5±2.3	27.3	81.1±5.3	23.1	104±7	9,173±471	87.9±7.2
40	1.13	0.035	310	1.50	31.4±1.8	19.1	51.4±3.6	23.7	75±5	6,733±205	89.5±7.0
42	1.45	0.034	240	1.61	43.1±4.2	23.2	67.0±5.8	23.5	91±7	8,460±337	93.4±8.2
Layer XIX											
45	1.005	0.065	770	2.46	46.2±3.7	21.1	68.7±5.5	22.8	92±7	9,056±449	98.8±8.9
47	2.17	0.072	790	2.38	95.9±4.7	38.4	136±9	23.4	159±10	13,12±885	82.4±7.7
49	1.091	0.084	870	2.55	52.3±4.1	23.3	76.4±6.2	23.4	100±7	8,478±368	84.9±7.3
77	1.149	0.342	1,610	1.95	44.8±3.6	30.9	77.0±5.9	23.1	100±7	9,615±443	95.9±8.1
Layer XXI											
1	1.675	0.045	440	1.69	52.6±5.7	28.2	81.8±7.5	21.8	104±8	11,41±455	109.9±9.9
2	2.205	0.087	370	1.52	62.5±7.9	35.5	99.9±9.9	21.6	122±11	10,84±516	89.2±8.9
61	1.451	0.049	840	1.58	42.6±4.6	28.2	71.8±6.3	21.8	94±7	8,518±436	90.9±8.7
Layer XXII											
65	2.234	0.05	450	1.97	81.5±4.4	36.4	119±8	23.4	143±9	12,36±693	86.6±7.4
66	1.138	0.022	220	1.92	40.4±4.4	18.5	59.4±5.7	23.5	83±7	7,573±340	91.2±8.7
67	1.065	0.054	190	1.97	39.1±2.1	17.2	56.9±4	23.5	80±6	6,879±262	85.4±6.9
Layer XXIII											
76	1.486	0.069	900	1.78	49.1±4.2	29.3	79.4±6.3	23.4	103±7	9,774±344	95.0±7.7

*The values of U, Th and K each have an error of ±6%. Fission-track analysis of five flints containing in excess of 1.5 p.p.m. uranium showed practically uniform distribution of the element.

†β dose equivalent to the flux of one α cm⁻², as deduced from fine grains measurements^{17,25}. The error on Sα ranges from ±5% to ±12%.

‡β dose equivalent to the α dose in flint calculated using the specific annual flux deduced from ref. 26.

§ The β dose in flints was calculated using the specific values given in ref. 27. The γ internal dose is less than 0.02 mGy yr⁻¹.

|| The external dose (estimated error: ±0.04 mGy) includes a cosmic contribution of 0.12 mGy yr⁻¹. It has been calculated using the data given in refs. 28 and 29. The thickness of overburden is about 3 m.

The uranium series radionuclides are essentially at equilibrium in the sediment: the activities of ²³⁸U, ²³⁴U, ²³⁰Th and ²¹⁰Po measured by α spectrometry are respectively 0.84±0.07, 0.78±0.06, 0.86±0.04 and 0.67±0.04 d.p.m. g⁻¹; the activities of ²²⁶Ra and ²¹⁴Bi measured by γ spectrometry, on another sample of sediment are respectively 0.70±0.08 and 0.71±0.02 d.p.m. g⁻¹.

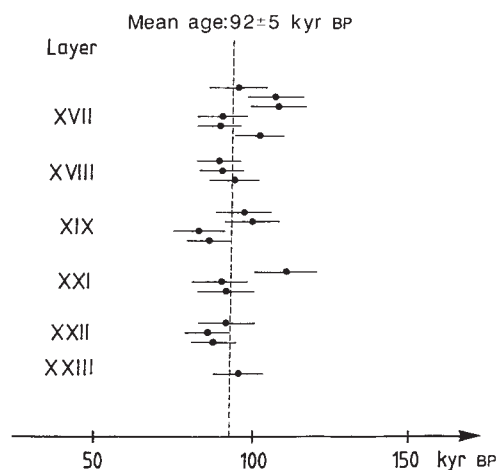


Fig. 1 Horizontal bars representing the thermoluminescence ages of burnt flints from Qafzeh cave as a function of depth of the unit in which the flints were found. The weighted average age deduced is given at the top. The errors (statistical plus systematic) at 68% confidence level are calculated as in ref. 30. A systematic error of ±7.5% has been assumed for the external dose to take into account possible variations in water content of the archaeological sediments (M. J. Aitken, personal communication).

(suggesting the opposite) can only be resolved by an independent dating technique, such as thermoluminescence (TL)^{6,16-18}.

Twenty burnt flints recovered from the Mousterian layers, which yielded Proto-Cro-Magnon remains 15 years ago, have been dated by TL. The experimental details of our dating method have been discussed elsewhere^{6,19,20}. Six specimens came from layer XVII, three from layer XVIII, four from layer XIX, three each from layers XXI and XXII and one from layer XXIII.

The external dose was measured using 13 dosimeters (CaSO₄/Dy), which were buried for a year within the sections surrounding the excavated deposits. As the majority of flints came from metre squares C11 and C12, the dosimeters were buried in adjacent squares B11, D11 and D12, within a distance of 1 m from the point of recovery of the flints.

The recorded external dose (γ plus cosmic) ranged from 0.22 to 0.27 mGy yr⁻¹, of which approximately 0.12 mGy yr⁻¹ was

due to cosmic rays. The concentrations of radioactive elements (^{238}U , ^{232}Th and ^{40}K) within the flints was determined by neutron activation analysis²¹. Due to the relatively low level of the external dose, the internal dose represents from ~63 to over 80% of the total annual dose received by the specimens. Therefore errors in the measured environmental dose have relatively little impact on the calculated ages (see Table 1).

In Fig. 1 the ages, with their statistical errors, are plotted against depth of flint recovery. These ages fall between 85 and 105 kyr and exhibit no systematic change with recovery depth, within the experimental errors. The age (weighted mean) deduced for layers XVII–XXIII is 92 ± 5 kyr (statistical + systematic error). Consequently, the TL data support the geological observation¹⁰ of a rather rapid accumulation of lower levels when the cave was occupied.

The TL dates lead to two main conclusions. First, archaic forms of modern humans classified as 'Proto-Cro-Magnons' were already present in southwest Asia during isotopic stage 5, ~92 kyr BP (ref. 22), and may have coexisted at a later date with a population of Neanderthals, one of whose remains has recently been dated at Kebara to ~60 kyr BP⁶. This makes the proposition of a close phylogenetic relationship between the two human types untenable and supports the contention that the Neanderthals were a stock of European origin and arrived relatively late in south west Asia. They may have migrated East when glacial conditions settled over large areas of the European continent^{1,23}. These west Asian dates should increase our knowledge of the place of origin and dispersal of modern man²⁴. Second, the stratigraphic sequence of Tabun is either of geographically limited value or of much greater antiquity than hitherto suspected²³. The Mousterian of Tabun D could still be older than the 'Tabun C' lithic industry which dominates the Qafzeh assemblages, if chronological reinterpretation of the Tabun cave place the 'Tabun D' type industry near the beginning of isotope stage 5 (ref. 22).

We thank the French Ministry of Foreign Affairs, the Centre National de la Recherche Scientifique, the Commissariat à l'Énergie Atomique and the L.S.B. Leakey Foundation for support.

Received 18 November 1987; accepted 15 January 1988.

- Vandermeersch, B. *Les hommes fossiles de Qafzeh, Israel* (CNRS, Paris, 1981).
- McCown, T.D. & Keith, A. *The stone Age of Mount Carmel* Vol. II (Clarendon, Oxford, 1939).
- Howells, W.W. *J. hum. Evolution* **5**, 477–496 (1976).
- Bar-Yosef, O. & Vandermeersch, B. In *Préhistoire du Levant* (eds Cauvin, J. & Sanlaville, P.) 282–286 (CNRS, Paris 1981).
- Vandermeersch, B. in *The Transition from the lower to the Middle Palaeolithic and the origin of modern Man* (eds Ronen, A.) 297–300 (Bar, Oxford, 1982).
- Valladas, H. *et al. Nature* **330**, 159–160 (1987).
- Jelinek, A. *J. Science* **216**, 4553, 1369–1375 (1982).
- Wolpoff, M. H. *Palaeoanthropology* (Knopf, New York 1980).
- Trinkaus, E. in *The Origins of Modern Humans* (eds Smith, F.H. & Spencer, F.) 251–293 (Liss, New York, 1984).
- Farrand, W.R. *J. arch. Sci.* **6**, 369–392, (1979).
- Tchernov, E. in *Préhistoire du Levant* (eds Cauvin, J. et Sanlaville, P.) 67–97 (CNRS, Paris, 1981).
- Haas, G. *Palaeovertebrata* **5**, 261–270 (1972).
- Tchernov, E. in *Quaternary extinctions* (eds Martin, P.S. & Klein, R.) 528–552 (University of Arizona Press, Tucson, 1984).
- Boutié, P. in *Investigations in South Levantine Prehistory* (eds Bar Yosef, O. & Vandermeersch, B.) (BAR, Oxford) (in the press).
- Copeland, L. in *Problems in Prehistory: Northeast Africa and the Levant* (eds Wendorf, F. & Marks, A.E.) 317–350 (SMU Press, Dallas, 1975).
- Aitken, M.J. *Thermoluminescence Dating* (Academic, New York, 1985).
- Huxtable, J. & Jacobi, R.M. *Archaeometry* **24**, 164–169 (1982).
- Bowman, S.G.E. & Sieveking, G. de G. *Pact* **9**, 353–361 (1982).
- Valladas, H., Geneste, J.M., Joron, J.L. & Chadelle, J.P. *Nature* **322**, 452–454, (1986).
- Valladas, H. & Valladas, G. *Archaeometry* **27**, 214–220 (1987).
- Chayla, B., Jaffrezic, H. & Joron, J.L. *C. r. hebdom. Acad. Sci., Paris* **277**, 273–275 (1973).
- Schackleton, N.J. & Opdyke, N.D. *Quat. Res.* **3**, 39–55 (1973).
- Bar-Yosef, O. in *Corridors, Cul de Sac and Coalescence* (ed. Trinkaus, E.) (Cambridge, University Press) (in the press).
- Gowlett, J.A.J. *Antiquity* **61**, 210–219 (1987).
- Valladas, H. & Valladas, G. *Pact* **6**, 171–178 (1982).
- Aitken, M.J. & Bowman, S.G.E. *Archaeometry* **17**, 132–138 (1975).
- Bell, W.T. *Archaeometry* **21**, 243–245 (1979).
- Kvasnicka, J. *Healths Phys.* **36**, 521–524 (1979).
- Prescott, J.R. & Stephan, L.G. *Pact* **6**, 17–25 (1982).
- Aitken, M.J. *Archaeometry* **18**, 233–238 (1976).

Pattern of covariation between life-history traits of European birds

Bernt-Erik Sæther

Game Research Division, Directorate for Nature Management, Tungasletta 2, N-7004 Trondheim, Norway

A large amount of variation is found in most reproductive traits of birds^{1,2}. Clutch size for instance, can vary from 1 to 15 between species of similar body weight. The adaptive significance of this variation is only poorly understood^{3,4}. According to life-history theory^{5,6}, large clutch size and early onset of reproduction are expected when the chances of survival are low. There is some support for the existence of such a relationship from studies of single species^{7–9}. Here I present evidence that, in European birds, clutch size is increased, and onset of reproduction occurs earlier in life, when the probability of survival is low.

This study is based on published adult survival rates of 107 species from 14 orders (data listed in ref. 10), making use of estimates based on ringing recoveries in which the number of returns is greater than 100. In most previous studies, mean adult survival rate has been computed using the estimators given by Lack¹¹ or Haldane¹². When no data on ringing recoveries were available, I included estimates based on the recovery rates of individually known birds, given that at least 100 birds have been at risk of dying. The two methods of estimating adult survival rate are strongly correlated¹⁰. Although several sources of error are involved in using those two methods^{13–15}, the precision of the estimates does reveal cross-species patterns. All variables were logarithmically transformed before analysis to obtain gaussian distributions.

The choice of taxonomic level is important when using the comparative method^{16–18}. Although analysis at lower taxonomic levels increases the sample size, such comparisons include data points that may not be statistically independent. A considerable proportion of the variance in clutch size (81.6%), age at maturity (36.6%) and adult survival rate (47.0%) was found at the level of family or above, using a nested analysis of variance^{19,20}. Here I analysed first for differences between families, and then between species within orders.

A close correlation was found across families between adult survival rate, clutch size and age at maturity (data on clutch size and age at maturity are listed for each genera in ref. 2). Clutch size decreased both with increasing age at maturity ($r = -0.74$, $n = 38$, $P < 0.001$) and adult survival rate ($r = -0.75$, $n = 38$, $P < 0.001$), as previously reported by O'Connor²¹. Maturation occurred at older ages in families in which the survival rate was high ($r = 0.83$, $n = 38$, $P < 0.001$)¹.

Body weight in birds is known to be correlated with age at maturity^{2,22}, clutch size²² and adult survival rate^{10,23}. Thus, the correlations between the life-history traits may simply be caused by their common correlation with body weight. To control for the effects of body weight I calculated the correlation between the residuals from the regression line of each life history trait on body weight. In families with a high adult survival rate for their body weight, clutch size was still relatively small (Fig. 1a, $r = -0.77$, $n = 38$, $P < 0.001$) and onset of reproduction occurred at relatively old age (Fig. 1b, $r = 0.75$, $n = 38$, $P < 0.001$). A similar pattern of covariation has previously been documented in mammals^{24–27}.

I also examined the patterns of covariation within orders. The relationships between clutch size, age at maturity and adult survival rates with body weight were established for the species within the orders Anseriformes, Charadriiformes and Passeriformes. These were the only orders in which estimates of adult survival rates of a sufficient number of species (>3) were available. I then computed the deviations of the species from the

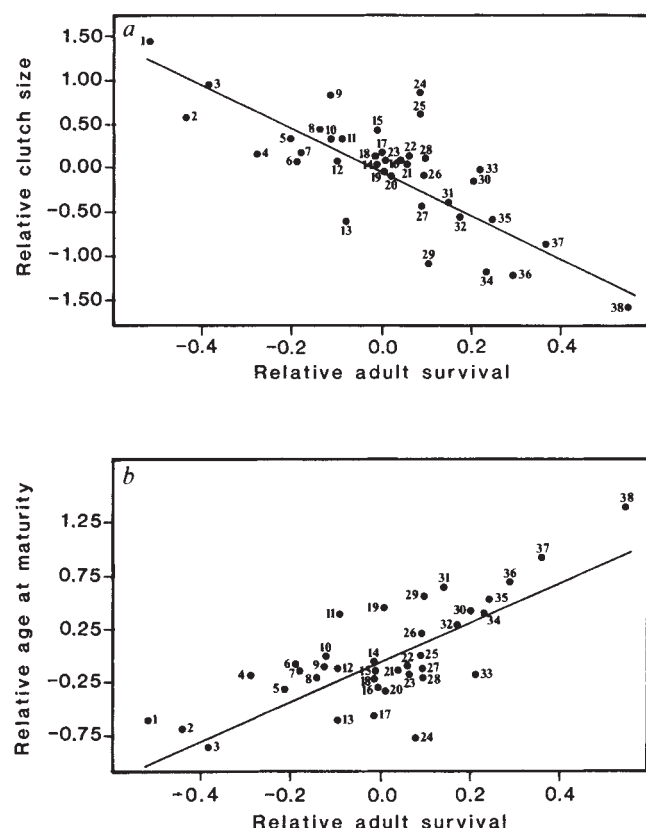


Fig. 1 Relative adult mortality rate in relation to (a) relative clutch size and (b) relative age at maturity of different families of European birds. The points represent the deviations of each family from the predicted value of the regression of the life history trait on body weight, using the family averages (logarithmically transformed) as data entries. 1, Phasianidae; 2, Tytonidae; 3, Tetraonidae; 4, Turdidae; 5, Sturnidae; 6, Hirundinidae; 7, Motacillidae; 8, Corvidae; 9, Anatidae; 10, Muscipapidae; 11, Ciconiidae; 12, Fringillidae; 13, Columbidae; 14, Sylviidae; 15, Ardeidae; 16, Scolopacidae; 17, Strigidae; 18, Emberizidae; 19, Phalacrocoracidae; 20, Charadriidae; 21, Accipitridae; 22, Passeridae; 23, Falconidae; 24, Rallidae; 25, Paridae; 26, Pandionidae; 27, Gaviidae; 28, Recurvirostridae; 29, Sulidae; 30, Haematopodidae; 31, Laridae; 32, Stercorariidae; 33, Alaudidae; 34, Alcidae; 35, Sternidae; 36, Procellariidae; 37, Apodidae and 38, Hydrobatidae.

value expected from these regressions, using separate regressions for each order. In all three orders maturation occurs relatively late when the adult survival rate is high ($r=0.65$, $n=18$, $P<0.01$; $r=0.65$, $n=25$, $P<0.001$ and $r=0.30$, $n=32$, $P>0.05$ for Anseriformes, Charadriiformes and Passeriformes, respectively). Also, Charadriiforme species with a high probability of survival for their body weight also laid relatively few eggs ($r=-0.68$, $n=25$, $P<0.001$), whereas no significant relationship appeared within the other two orders.

Elsewhere, I have shown that juvenile and adult survival rate is positively correlated¹⁰, that is, the life expectancy of an offspring at independence is higher in some species than in others. The results presented here show that reproductive rates change as a response to variation in survival rates. This conclusion agrees with the results from several models of evolution of optimal life histories²⁸⁻³⁰, which predict that a small increase in reproductive success later in life either by reducing clutch size or by delaying onset of reproduction is more likely to be favoured when the chances of future survival is high. A similar relationship between adult life expectancy and age at maturity, even after controlling for the effects of body weight, has pre-

viously been described in mammals^{24,25}. Similarly, longevity is inversely related to clutch size in one group of North-American birds when the effects of body weight are controlled with partial correlation analysis³¹. Alternatively, the present pattern could be the result of increased density-dependent mortality in species with relative large clutch size. Until more data are available on the effects of density-dependent processes on mortality rates in different species of birds, we cannot reject any of these hypotheses.

I thank P. Harvey, M. Elgar, M. Pagel, D. Promislow, A. Grafen and A. Harper for help with the analyses or comments on the manuscript.

Received 14 September; accepted 18 December 1987.

1. Lack, D. *Ecological Adaptations for Breeding in Birds* (Methuen, London, 1968).
2. Sæther, B.-E. *Oikos* **48**, 79-88 (1987).
3. Winkler, D. W. & Walters, J. R. *Curr. Ornithol.* **1**, 33-68 (1983).
4. Murphy, E. C. & Haukioja, E. *Curr. Ornithol.* **4**, 141-180 (1986).
5. Williams, G. C. *Adaptation and natural selection* (Princeton University Press, 1966).
6. Williams, G. C. *Am. Nat.* **100**, 687-690 (1966).
7. Bell, G. *Evolution* **38**, 300-313 (1984).
8. Reznick, D. *Oikos* **44**, 257-267 (1985).
9. Bryant, D. M. *J. Anim. Ecol.* **48**, 655-675 (1979).
10. Sæther, B.-E. *Ornis Scand.* (in the press).
11. Lack, D. *Proc. Orn. Congr.* **10**, 409-448 (1951).
12. Haldane, J. B. S. *Acta XI Congr. Int. Orn.* **1954**, 454-458 (1955).
13. Cave, A. J. *Vogelwarte Sonderheft* **29**, 160-171 (1977).
14. Lakhani, K. H. & Newton, I. *J. Anim. Ecol.* **52**, 83-91 (1983).
15. Anderson, D. R., Burnham, K. P. & White, G. C. *J. Anim. Ecol.* **54**, 89-98 (1985).
16. Harvey, P. H. & Mace, G. M. in *Current Problems in Sociobiology* (eds Kings College Sociobiology Group) 343-361 (Cambridge University Press, Cambridge, 1982).
17. Felsenstein, J. *Am. Nat.* **125**, 1-15 (1985).
18. Martin, R. D. & Harvey, P. H. in *Size and scaling in primate biology* (ed. Jungers, W. L.) 147-173 (Plenum, New York, 1985).
19. Harvey, P. H. & Clutton-Brock, T. H. *Evolution* **39**, 559-581 (1985).
20. Sokal, R. R. & Rohlf, F. J. *Biometry* 2nd edn (Freeman, New York, 1981).
21. O'Connor, R. J. in *Behavioural Ecology, Ecological Consequences of Adaptive Behaviour* (eds Sibley, R. M. & Smith, R. H.) 105-142 (Blackwell, Oxford, 1985).
22. Western, D. & Semakula, J. *Oecologia* **54**, 281-290 (1982).
23. Calder, W. J. III *J. theor. Biol.* **102**, 135-144 (1983).
24. Harvey, P. H. & Zammuto, R. M. *Nature* **315**, 319-320 (1985).
25. Sutherland, W. J., Grafen, A. & Harvey, P. H. *Nature* **320**, 88 (1986).
26. Millar, J. S. & Zammuto, R. M. *Ecology* **64**, 631-635 (1983).
27. Stearns, S. C. *Oikos* **41**, 173-187 (1983).
28. Michod, R. E. *Am. Nat.* **113**, 531-550 (1979).
29. Roughgarden, J. *Theory of population genetics and evolutionary ecology: an introduction* (Macmillan, New York, 1979).
30. Charlesworth, B. *Evolution in age-structured populations* (Cambridge University Press, Cambridge, 1980).
31. Zammuto, R. M. *Can. J. Zool.* **64**, 2739-2749 (1986).

Evidence for a spectral basis of texture perception in bat sonar

Sabine Schmidt

Zoologisches Institut der Universität, Luisenstrasse 14,
D-8000 München 2, FRG

Bats obtain information about the structure of objects in the outside world from their echolocation signals¹⁻⁴, an extremely useful method when hunting non-flying prey in densely cluttered habitats, for example. Information about object structure is contained both in the time and in the spectral interference patterns of signals reflected from surfaces at different distances from the bat. I report here an experiment designed to test the extent to which bats use these two types of information. A 'phantom target' is generated by playing back to an echolocating bat signals that mimic the result of reflection from two planes set at different distances. The ability of the bat to discriminate between two such targets is investigated as a function of the separations of the planes. Several of the results do not fit the hypothesis that the bat simply uses time-delay information: the very small time difference that can be discriminated, the fall off in ability to discriminate planes at a particular separation and the symmetry of the discrimination ability measured in the frequency domain. The empirical data can best be fitted by a function based on spectral correlation.

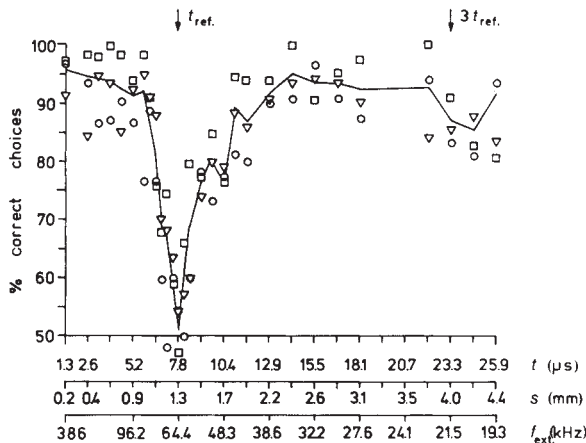


Fig. 1 Performance of three bats in choosing a reference target with $t_{\text{ref}} = 7.77 \mu\text{s}$ internal delay as a function of internal delays (t), in terms of the distance between two parallel planes (s) and of the position of the first spectral notches (f_{ext}) in the echo spectrum; note that data are plotted on a linear time scale (t). In contrast, a plot on a linear frequency scale (f_{ext}) would show a symmetrically shaped valley at $f_{\text{ref}} = 64.4 \text{ kHz}$. Data points result from more than 30 trials; so 75% correct choices are significantly different from random choice ($P < 0.01$).

Megaderma lyra, a gleaning bat feeding on insects and small vertebrates, emits short ($\sim 0.5 \text{ ms}$), low intensity (75–80 dB SPL (sound pressure level relative to $2 \times 10^{-5} \text{ N m}^{-2}$)), broadband (20–120 kHz) sonar pulses containing several harmonics. Three individuals were successfully trained to discriminate between phantom targets offered in a forced two-choice behavioural test. Vocalizations of the bats hanging at a starting position were recorded using a QMC Instruments microphone mounted 13 cm below the bat's head and fed into a phantom-target generator. Two-front phantom targets were produced on-line by adding two replicas of the bat's vocalizations with different delays, T , and $T + t_i$. The absolute delay, T , was chosen such that phantom targets appeared at a distance of 1.34 m from the bats, and 33 cm in front of the ultrasonic speakers to avoid overlap between phantom echoes and clutter from the setup. Echo structure is determined by t_i only, the internal delay between the two replicas, which could be varied in steps of 81 ns. In the frequency domain, the interference patterns corresponding to the t_i chosen were checked by measuring the frequency response of the phantom-target generator. The resultant attenuation peaks (spectral notches) in the echo spectra exceeded 40 dB. Phantom echoes were played back through two ultrasonic speakers with flat frequency response. Maximum echo amplitudes of $\sim 80 \text{ dB SPL}$ could be achieved.

The bats were initially trained to select a rewarded reference target (internal delay $t_i = t_{\text{ref}} = 7.77 \mu\text{s}$, equivalent to a spectral notch at a frequency, f_{ref} , of 64.4 kHz and a target depth of 1.3 mm) from a target with different t_i . They could freely probe each of the two targets offered in a trial by directing their calls towards the corresponding speaker, which was electronically selected for phantom-target output. Bats indicated their choices by flying directly to one of the two feeding places associated with the targets. Correct responses were rewarded with mealworms. In the experimental trials, target presentation was randomized. Up to nine different target pairs were offered in one experimental session. The target combinations presented were unknown to the experimenter during the session. Occasionally target pairs not including the reference were offered. Data were pooled from several days.

The performance of the bats for trials in which the reference target was presented is given in Fig. 1. Three features are remarkable. First, whereas the bats choose at random for $t_i = t_{\text{ref}}$, an arbitrary 75% threshold criterion is met for target pairs differing

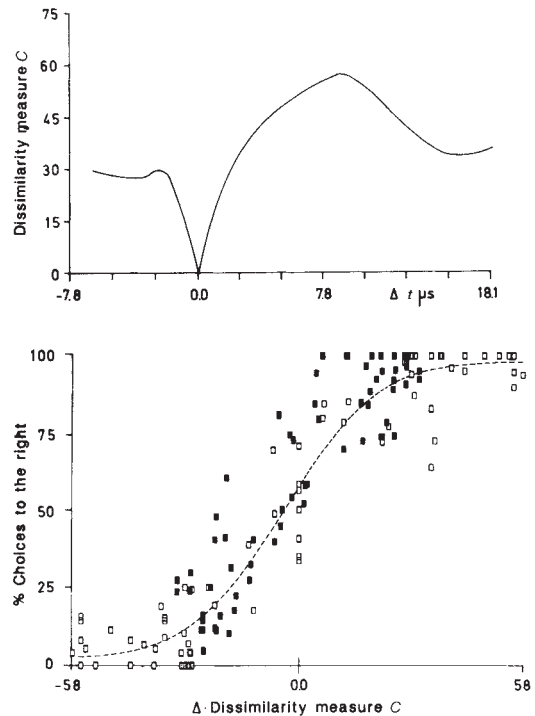


Fig. 2 *a*, Dissimilarity measures for the targets calculated by spectral correlation. C is given as a function of differences in internal delay relative to the reference target:

$$C := \arccos \left(\frac{\int_{f_{\text{r. range}}} b(f) \mathbf{R}(t_{\text{ref}}, f) \mathbf{R}(t_i, f) df}{\left(\int_{f_{\text{r. range}}} b(f) \mathbf{R}(t_{\text{ref}}, f)^2 df \right)^{1/2} \left(\int_{f_{\text{r. range}}} b(f) \mathbf{R}(t_i, f)^2 df \right)^{1/2}} \right)$$

$$\mathbf{R}(t_i, f) := (1 + \cos 2\pi f t_i)^{1/2}$$

Definitions and parameters: $b(f)$: spectral weighting function, t_i : internal delay, fr. range = 15–100 kHz, $t_{\text{ref}} = 7.77 \mu\text{s}$,

$$b(f) = \begin{cases} 1 & \text{for } 25 \text{ kHz} \leq f \leq 90 \text{ kHz} \\ 0 & \text{for } f < 15 \text{ kHz or } f > 100 \text{ kHz} \end{cases}$$

sigmoid transition: 15–25 kHz, 90–100 kHz

C represents the angle between the two normalized vectors $\mathbf{R}(t_{\text{ref}}, f)$ and $\mathbf{R}(t_i, f)$ in the vector space of frequency-dependent functions. For $\mathbf{R}(t_{\text{ref}}, f) = \mathbf{R}(t_i, f)$ the two vectors coincide and dissimilarity equals zero. $\mathbf{R}(t_i, f)$ gives the echo spectrum returning from a target with the internal delay t_i for an idealized broadband call with flat spectrum. *b*, Performance data for one bat (% choices to the right) plotted as a function of the differences between the dissimilarity measures obtained for the target pairs offered. Open squares represent data points from target combinations including, filled squares from those not including the reference. As expectation values even for simple depth discrimination tasks are less than 100%, modified Gaussian error functions, asymptotic at 2% and 98%, are used for approximation. All data points are well represented by the function, although only those from target combinations including the reference were approximated.

only $\sim 1 \mu\text{s}$ in internal delay, equivalent to a depth difference of 0.2 mm. Second, performance is reduced for internal delays near $t_i = 3t_{\text{ref}}$, where the spectral notches of the target pairs offered are harmonically related. Third, performance is symmetrical about the reference target in the frequency domain. Due to the inverse relationship between time and frequency, symmetry is distorted for data shown as a function of time. In a more general situation, target pairs not including the reference were offered; the bats still continued to choose between targets, a phenomenon that must be accounted for by the model framework proposed.

It is assumed that the bats acquire a memory representation of the rewarded reference target during the training procedure. In the experimental trials, the target more similar to this standard

is chosen. The problem, then, is to find the criteria of similarity used. A model in the proposed framework consists of (1) a function that assigns to a given target a number characteristic of its dissimilarity with the reference, and (2) a second function that predicts the bat's behaviour from these dissimilarity measures for the two targets offered in a trial. Decisions are expressed as a function of the difference in those measures and approximated by a gaussian error function. Dissimilarity measures supposing time or else frequency processing by the animals were tested. In the frequency-domain models, the overlap of the reflection spectra of the reference and a given target determined in different ways was used. All these frequency-domain interpretations predict the reduced performance for $t_i = 3t_{ref}$, and a spectral symmetry in the performance curve. The demanded frequency resolution is in the kHz range, a requirement supported by auditory physiology. The model based on spectral correlation gave the best approximations (according to χ^2 test) of the bats' performance (see Fig. 2).

Time-domain models assume that time differences of the order of $1\ \mu\text{s}$ are determined. This level of time-difference determination has not yet been shown in mammalian physiology.

When time differences were used as a dissimilarity measure, data points systematically deviated from the approximation in all three bats. Although this cannot justify a general refutation of time-domain models, I suggest that the unified description of echo processing in bats in the time domain⁵ should be abandoned in favour of a more pragmatic one⁶ explaining texture discrimination in spectral terms. Then, texture perception by bats, rather than being a specific adaptation, reveals more about how diverse properties represented in complex spectral patterns may form an image in the auditory areas of the brain.

This work was supported by SFB 204 München. I thank Professor G. Neuweiler and M. Betz.

Received 21 September 1987; accepted 6 January 1988.

1. Griffin, D. R., Friend, J. H. & Webster, F. A. *J. exp. Zool.* **158**, 155–168 (1965).
2. Bradbury, J. W. *J. exp. Zool.* **173**, 23–46 (1970).
3. Simmons, J. A. *et al. Science* **186**, 1130–1132 (1974).
4. Habersetzer, J. & Vogler, B. *J. comp. Physiol.* **152**, 275–282 (1983).
5. Simmons, J. A. & Stein, R. A. *J. comp. Physiol.* **135**, 61–84 (1980).
6. Beuter, K. J. in *Animal Sonar Systems* (eds Busnel, R.-G. & Fish, J. F.) 747–761 (Plenum, New York, 1980).

Comparison of simian immunodeficiency virus isolates

Harry W. Kestler III, Yen Li, Yathirajulu M. Naidu, Carole V. Butler, Michael F. Ochs, Gerlinde Jaenel*, Norval W. King, Muthiah D. Daniel & Ronald C. Desrosiers†

New England Regional Primate Research Center, Harvard Medical School, 1 Pine Hill Drive, Southborough, Massachusetts 01772, USA

Information on the extent of genetic variability among non-human primate lentiviruses related to human immunodeficiency virus (HIV) is sorely lacking. Here we describe the isolation of two molecular clones from the simian immunodeficiency virus (SIV) and their use to derive restriction endonuclease maps of five SIV isolates from rhesus macaques and one from a cynomolgus macaque. Although similar, all six viral isolates are readily distinguishable; the single isolate from a cynomolgus macaque is the most different. The restriction endonuclease map of one macaque isolate (SIV_{MAC-251}) is identical to that published by others for STLV-III_{AGM} of African green monkeys^{1,2} and for HTLV-IV of humans³. Nucleotide sequences from the envelope region of cloned SIV_{MAC-251} have more than 99% identity to previously published sequences for STLV-III_{AGM} (refs 2, 4) and HTLV-IV (ref. 4). These results and other observations provide strong evidence that isolates previously referred to as STLV-III_{AGM} and HTLV-IV by others are not authentic, but were derived from cell cultures infected with SIV_{MAC-251}.

Quantitation of genetic relatedness among human and non-human primate lentiviruses should help elucidate the history and evolution of this group of viruses and the origins of viruses that cause AIDS in humans. Details of the genetic composition of the non-human primate lentiviruses will be useful in animal models for studying AIDS and can be exploited to understand the molecular basis for cellular tropism, species specificity and pathogenicity.

SIV was first isolated from captive rhesus macaques (*Macaca mulatta*) at the New England Regional Primate Research Center

(NERPRC)⁵. Macaques are an Asian Old World primate. SIV-reactive antibodies were subsequently demonstrated in 30–50% of green monkeys (*Cercopithecus aethiops*) from central Africa⁶ and this was soon followed by a report of isolation of an SIV called STLV-III_{AGM} from African green monkeys by Kanki *et al.*⁷. SIV has also been isolated from sooty mangabeys^{8–10} (an African primate), from a pig-tailed macaque (*M. nemestrina*)¹¹, from a cynomolgus macaque (*M. fascicularis*) (see below), and more recently from African green monkeys (ref. 12, and see below). Recent serological surveys suggest that some other species of African primates may also be infected with SIV- or HIV-related viruses^{10,12}. Lentiviruses have been isolated from humans in west Africa that are closer to SIV_{MAC} than to HIV-1, the AIDS virus prevalent in the United States, Europe and Central Africa. These isolates from humans in West Africa were originally called HTLV-IV by Kanki, Essex and colleagues¹³ and LAV-2 by Clavel *et al.*¹⁴, but are now generally referred to as HIV-2. As described below, the initial STLV-III_{AGM} (ref. 7) and HTLV-IV (ref. 13) isolates have apparently been derived from SIV_{MAC-251}-infected cell cultures.

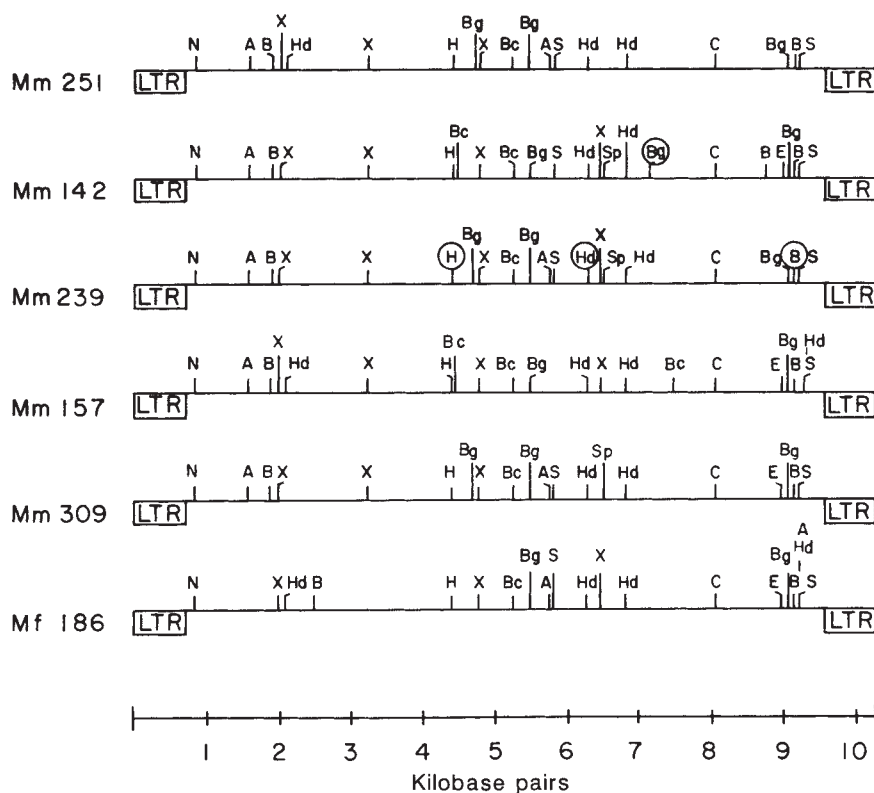
The SIV_{MAC} isolates used in this study were all from animals housed at the NERPRC. Isolates 251, 239 and 142 from individual *M. mulatta* have been described previously⁵. (Viral isolates are designated with the same number as the animal from which they were derived.) Isolates 157 and 309 were also from individual *M. mulatta* and 186 was from a *M. fascicularis* identified in our 1986 serological survey¹⁵. Two pairs of the isolates are known to be related by *in vivo* transmission. In one case, minced SIV-containing lymphoma tissue from 251 was inoculated into other macaques; following two additional successive *in vivo* tissue transmissions, SIV was recovered from a final recipient 239 (ref. 5). SIV was present *in vivo* for a total of 251 days between the times that samples were taken for SIV isolation. Macaque 142 became infected *in utero* from his mother 157 (ref. 15); SIV was recovered from 142 at 18 months of age and SIV was recovered from 157 34 months after she had transmitted SIV to her offspring. Macaques 157, 309 and 186 were never inoculated with biological material and became infected naturally while at NERPRC. In the studies described below, HUT-78 cells infected with uncloned, non-plaque-purified SIV were used.

Using bacteriophage lambda EMBL4 as vector, full-length molecular clones containing integrated proviral DNA from SIV_{MAC-251} and SIV_{MAC-239}-infected HUT-78 cells were obtained (Fig. 1). Chakrabarti *et al.* have previously described the isolation of a full-length molecular clone of SIV_{MAC-142} (ref.

* Present address: Institute für Klinische Virologie, Universität Erlangen-Nürnberg, 852 Erlangen, Loschgestrasse 7, FRG.

† To whom reprint requests should be addressed.

Fig. 1 Restriction endonuclease maps of SIV_{MAC} DNA present in infected cells. To isolate molecular clones, total cell DNA was first prepared from HUT-78 cells infected with SIV_{MAC-251} and SIV_{MAC-239}. EcoRI-digested infected-cell DNA was used to prepare libraries in lambda EMBL4. Using the 7.3 kilobase fragment from pK2-10BamA described by Hirsch *et al.*¹ as a probe, full-length clones for SIV_{MAC-251} and for SIV_{MAC-239} were isolated for further analysis, including restriction endonuclease mapping. A full-length clone for SIV_{MAC-142} has been described previously¹⁶. For analysis of viral DNA present in Hirt supernatants, HUT-78 cells were infected with each of the six SIV isolates shown and replicative intermediate DNA was prepared 48 h later as previously described³². Restriction maps were constructed based on the results of single and double digestions, gel electrophoresis, and Southern transfer hybridization. Bg (ringed) indicates a BglII site present in the infectious molecular clone of 142 that was not detected in SIV_{MAC-142}-infected cell, Hirt supernatant DNA and was not present in other independent 142 clones. Hd (ringed) similarly indicates a HindIII site present in the infectious clone of 239 that was not detected in 239-infected cell, Hirt supernatant DNA. B and H (ringed) are BamHI and HpaI sites absent in the infectious molecular clone of 239 that were present in 239-infected cell, Hirt supernatant DNA. N, NarI; A, AvaI; B, BamHI; X, XbaI; Hd, HindIII; H, HpaI; Bg, BglII; Bc, BclI; S, SstI; Sp, SphI; C, ClaI.



16). All three of these cloned SIV_{MAC} DNAs are infectious upon transfection into HUT-78 cells using DEAE-dextran (our unpublished observations). Restriction endonuclease maps were derived for lambda 251 and lambda 239 cloned DNA by analysis of a series of single, double and triple digests.

Replicative intermediate Hirt supernatant DNA was prepared from HUT-78 cells infected with each of the six isolates described above (Fig. 1). DNA was digested with restriction endonucleases (various combinations of single and double digests), electrophoresed through agarose gels, Southern transferred and hybridized with cloned SIV_{MAC} DNA. The resultant restriction endonuclease maps showed that all six isolates were readily distinguishable (Fig. 1). Minor viral DNA species were not detected in the Hirt supernatant DNA, for example with 50% of the molecules having a particular restriction endonuclease site and 50% lacking it. The map of viral DNA in SIV_{MAC-251}-infected cells was identical to that of the isolated lambda clone. In other cases, differences were noted. The SIV_{MAC-142} infectious molecular clone has a BglII site (circled in Fig. 1) which was not detected in the infected cell DNA and was not present in other SIV_{MAC-142} clones that we have independently isolated. The infectious molecular clone of SIV_{MAC-239} actually has three differences when compared to viral DNA present in the Hirt supernatant of SIV_{MAC-239}-infected cells (Fig. 1).

We have tabulated the restriction-endonuclease-site conservation of the viral DNAs present in the Hirt supernatant of infected cells (Table 1). The conservation between isolates varied from 86% (18 of 21 sites conserved in isolate 251 versus 309) to 56% (only 14 of 25 sites conserved in isolate 186 versus 239). 86% and 56% conservation of 6-cut restriction sites correspond to ~98% and 93% sequence conservation. Inspection of the restriction endonuclease maps shows that the SIV isolate from the cynomolgus macaque 186 stands out as the most different. Even the isolates related through *in vivo* transmission were readily distinguishable; isolates 251 and 239 had 17 of 21 (81%) sites conserved and isolates 157 and 142 had 18 of 24 (75%) sites

conserved. It is not presently known whether a mix of SIV with variable restriction maps was transferred into the recipient animal, or if variations emerged with time following the *in vivo* transmission, or both. The situation should become clearer as we follow macaques experimentally infected with SIV derived from infectious molecular clones.

In contrast to this variability among SIV_{MAC} isolates, the restriction endonuclease map of SIV_{MAC-251} was identical to that published for STLV-III_{AGM} (refs 1, 2) and for HTLV-IV (ref. 3). To investigate this unusual similarity, random M13 subclones from the *env* region of the SIV_{MAC-251} infectious genomic clone were sequenced using the dideoxy chain termination method. For each region, the sequences of both DNA strands were determined. Of 1,035 nucleotides sequenced in *env* of SIV_{MAC-251}, only 7 differences were noted with one published sequence for STLV-III_{AGM} (ref. 2) and only 9 differences were noted with other published sequences for STLV-III_{AGM} and HTLV-IV (ref. 4); of 1,035 nucleotides sequenced in *env* of SIV_{MAC-251}, 34 differences were noted with the published sequence for SIV_{MAC-142} (ref. 16). Thus 96.7% sequence conservation was measured within *env* of the two SIV_{MAC} isolates 251 and 142, and more than 99% conservation was observed when SIV_{MAC-251} was compared with the published sequences for STLV-III_{AGM} and HTLV-IV. SIV_{MAC-251} sequences corre-

Table 1 Restriction endonuclease site conservation among SIV_{MAC} isolates

	239	142	157	186	309
251	17/21	16/24	16/24	15/24	18/21
239		17/23	15/25	14/25	18/21
142			18/24	15/26	18/23
157				16/25	16/25
186					15/25

The restriction endonuclease maps were derived from viral DNA present in the Hirt supernatant of infected cells (see Fig. 1).

sponding to nucleotide numbers 6,906–7,239, 7,257–7,374, 7,385–7,774 and 7,855–8,047 (ref. 2) were sequenced. When compared to the sequence of Franchini *et al.*², only 7 differences were noted at nucleotide numbers 7,170 (G to A), 7,210 (A to G), 7,443 (G to A), 7,475 (T to C), 7,535 (G to A), 7,931 (C to T) and 7,934 (T to C).

The following can be cited as evidence that isolates called STLV-III_{AGM} and HTLV-IV were actually derived from SIV_{MAC-251}-infected cell cultures. (1) Five original isolates of STLV-III_{AGM} (ref. 7) did not vary as expected when restriction endonuclease maps were compared³. (2) Two isolates of HTLV-IV (ref. 13) did not vary as expected when restriction endonuclease maps were compared³. Furthermore, the restriction maps of these HTLV-IV were identical to those of STLV-III_{AGM} and 99% conservation was observed in the regions that were sequenced^{2–4}. Such conservation would be highly unusual for authentic independent isolates. Not only do independent isolates of HIV or SIV vary from one another, but even independent isolates of HIV-1 from the same individual can vary¹⁷. In this report, we have shown that successive SIV_{MAC} isolates following *in vivo* transmission can differ in their restriction endonuclease maps. (3) Although six isolates of SIV_{MAC} were readily distinguished here, one of them, SIV_{MAC-251}, had the same restriction endonuclease map as STLV-III_{AGM} and HTLV-IV and had more than 99% sequence identity within *env* when compared to published sequences for STLV-III_{AGM} and HTLV-IV. (4) SIV_{MAC-251} was one of two isolates originally provided to Drs P. Kanki and M. Essex in October–November 1984 (ref. 18). (5) SIV_{MAC-251} was the prototype used in their early studies¹⁹. (6) The nucleotide sequence of HIV-2 isolated by Clavel *et al.*^{14,20} has only about 75% nucleotide conservation when compared to SIV_{MAC} (ref. 16). Restriction-site conservation between HIV-2_{ROD} and SIV_{MAC-142} is less than 20%. Isolates of HIV-2 by Clavel *et al.* from humans in west Africa do exhibit intratypic variation²¹. Furthermore, a virus isolate of Hahn *et al.*²² from the same group of prostitutes used by Kanki and Essex resembles LAV-2 (HIV-2), not HTLV-IV. (7) Ohta *et al.* have recently described the isolation and characterization of SIV_{AGM} (ref. 12). Their SIV_{AGM} differs significantly from SIV_{MAC}, forming a distinct, related virus grouping^{12,23}. Using procedures developed by Ohta *et al.*¹², Daniel *et al.* have also succeeded in isolating SIV_{AGM} (manuscript in preparation); this SIV_{AGM} is only distantly related to SIV_{MAC}, as demonstrated by the low stringency (41 °C below *T_m* of homologous hybrid) necessary to detect cross-hybridizing DNA sequences and the dissimilarity of their restriction endonuclease fragmentation profiles.

The restriction enzyme maps of the different SIV_{MAC} isolates as present in the infected cell cultures are readily distinguishable, with only microheterogeneity within each isolate. SIV_{MAC-251}, STLV-III_{AGM} and HTLV-IV are the only isolates that are so precisely alike. The evidence that STLV-III_{AGM} and HTLV-IV were derived from SIV_{MAC-251}-infected cell cultures is now overwhelming. The number of publications which purport to describe STLV-III_{AGM} and HTLV-IV and which are thus affected by these observations include refs 1–4, 7, 13, 24–31.

At this time we do not know the origin of SIV in our macaque monkey colony. Infection of macaques in captivity with SIV is rare. Only 3 macaques out of 848 were positive for SIV antibody in our 1986 survey, and we have been able to document a total of only 11 natural (that is, not from virus or tissue inoculation) SIV infections in NERPRC macaques (ref. 15). The SIV isolates from rhesus macaques in this study are sufficiently similar that a single point source could possibly have been the origin of most or all NERPRC rhesus macaque infections, but we do not know if this is actually so. It is even possible that an animal of some other genus could have been a source for NERPRC macaque infection. It should be pointed out, however, that the SIV_{AGM} and SIV mangabey isolates that have been examined to date are so different from SIV_{MAC} isolates at NERPRC that such viruses would not be a likely source for NERPRC macaque

infection. Much more work is needed to detail the genetic make-up of SIV isolates from a variety of feral as well as captive Old World primates.

We thank Dr James Mullins for providing the pK2-10BamA SIV clone before publication, Dr Masanori Hayami for informing us of his results before publication, Diane Schmidt and Cynthia Finger for technical assistance and Joanne Newton, Nancy Adams and Beverly Blake for preparing and editing the manuscript. This work was supported by PHS grants from the NIH and by the Milton Fund and the Whitaker Health Sciences Fund. Research was also sponsored in part by the NCI, DHHS, under subcontract with Program Resources Incorporated. The sequence data in this publication have been submitted to the EMBL/GenBank Libraries under accession number Y00294.

Received 19 October; accepted 23 December 1987.

1. Hirsch, V. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 9754–9758 (1986).
2. Franchini, G. *et al. Nature* **328**, 539–543 (1987).
3. Kornfeld, H., Riedel, N., Viglianti, G. A., Hirsch, V. & Mullins, J. I. *Nature* **326**, 610–613 (1987).
4. Hirsch, V., Riedel, N. & Mullins, J. I. *Cell* **49**, 307–319 (1987).
5. Daniel, M. D. *et al. Science* **228**, 1201–1204 (1985).
6. Kanki, P. J. *et al. Lancet* **i**, 1330–1332 (1985).
7. Kanki, P. J., Alroy, J. & Essex, M. *Science* **230**, 951–954 (1985).
8. Fultz, P. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 5286–5290 (1986).
9. Murphy-Corb, M. *et al. Nature* **321**, 435–437 (1986).
10. Lowenstine, L. J. *et al. Int. J. Cancer* **38**, 563–574 (1986).
11. Benveniste, R. E. *et al. J. Virol.* **60**, 483–490 (1986).
12. Ohta, Y. *et al. Int. J. Cancer* **41**, 115–122 (1988).
13. Kanki, P. J. *et al. Science* **232**, 238–243 (1986).
14. Clavel, F. *et al. Science* **233**, 343–346 (1986).
15. Daniel, M. D. *et al. Int. J. Cancer* (in the press).
16. Chakrabarti, L. *et al. Nature* **328**, 543–547 (1987).
17. Hahn, B. H. *et al. Science* **232**, 1548–1553 (1986).
18. Desrosiers, R. C., Daniel, M. D., Letvin, N. L., King, N. W. & Hunt, R. D. *Nature* **327**, 107 (1987).
19. Kanki, P. J. *et al. Science* **228**, 1199–1201 (1985).
20. Guyader, M. *et al. Nature* **326**, 662–669 (1987).
21. Clavel, F. *et al. Nature* **324**, 691–695 (1986).
22. Hahn, B. H. *et al. Nature* **330**, 184–186 (1987).
23. Schneider, J. *et al. Annls Inst. Pasteur Virol.* **138**, 93–99 (1987).
24. Denis, F. *et al. Lancet* **i**, 408–411 (1987).
25. Kanki, P. J. *et al. Science* **236**, 827–831 (1987).
26. Arya, S. K. *et al. Nature* **328**, 548–550 (1987).
27. Norrby, E. *et al. Nature* **329**, 248–250 (1987).
28. Hendry, R. M. *et al. Lancet* **ii**, 455 (1986).
29. Guo, H.-G. *et al. AIDS Res. Hum. Retroviruses* **3**, 177–185 (1987).
30. Franchini, G. *et al. AIDS Res. Hum. Retroviruses* **3**, 11–17 (1987).
31. Albert, J. *et al. AIDS Res. Hum. Retroviruses* **3**, 3–10 (1987).
32. Desrosiers, R. C. *et al. J. Virol.* **54**, 552–560 (1985).

ESSEX AND KANKI REPLY—Recently, certain data^{1–3} and comment^{4,5} has centred around the authenticity of certain HIV-2/HTLV-IV and STLV-III_{AGM} isolates. We acknowledge the results of others and agree that certain isolates initially reported by us as HIV-2/HTLV-IV should be considered SIV unless proven otherwise. At present our ongoing studies use other independent isolates of HIV-2.

The antigenic relatedness of both SIV and HIV-2 to the prototype HIV-1 virus prompted the discovery and further categorization of these related viruses^{6–8}. The conclusions drawn from these studies remains unchanged, irrespective of the authenticity of certain isolates. We first recognized the existence of SIV_{MAC} as a primate relative of HIV-1. This work was done in collaboration with workers at the NERPRC, the serologic analyses were conducted in our laboratory^{6,9}. We first reported the existence of SIV infection in significant numbers of certain wild African primate species, again based on the presence of SIV-reactive antibodies^{7,10}.

As Kestler *et al.*³ state, there is still a lack of understanding of which primates species are naturally infected with SIVs and where they originated from. What is the origin of the SIV_{MAC} virus? Epidemiological studies have demonstrated a surprising absence of SIV-infected macaques in the wild^{11,12}. One obvious possibility is that the macaque is a susceptible but unnatural host which became infected with SIV from another primate species in captivity.

In 1985 we published the first evidence of an HIV-2 type of infection in people⁸. Healthy prostitutes in West Africa were found to have antibody profiles that were indistinguishable from

those found in SIV-infected monkeys^{8,13}. Descriptions of HIV-2 that followed by others were also based on the detection of antibodies in people that were highly cross-reactive with SIV^{14,15}.

Sero-epidemiological analyses done in our laboratory, and confirmed elsewhere, have established that SIV and HIV-2 can be used interchangeably as an antigen source, whether screening for antibodies in monkeys or people^{8,14-16}. All conclusions published by other laboratories as well as our own concerning the epidemiological distribution and natural biology of SIV and HIV-2 remain as originally stated. A better understanding of both SIV and HIV-2 will be gained by further studies on the biology of these viruses in their naturally infected hosts.

M. ESSEX
P. KANKI

Department of Cancer Biology,
Harvard School of Public Health,
Boston, Massachusetts 02115, USA

- Kornfeld, H. *et al. Nature* **326**, 610-613 (1987).
- Hahn, B. *et al. Nature* **330**, 184-188 (1987).
- Kestler, H. *et al. Nature* (this issue).
- Newmark, P. *Nature* **327**, 458 (1987).
- Desrosiers, R. *et al. Nature* **327**, 107 (1987).
- Kanki, P. *et al. Science* **228**, 1199-1201 (1985).
- Kanki, P. *et al. Lancet* **i**, 1330-1332 (1985).
- Barin, F. *et al. Lancet* **ii**, 1387-1389 (1985).
- Daniel, M. *et al. Science* **228**, 1201-1204 (1985).
- Kanki, P. *et al. Science* **230**, 951-954 (1985).
- Kanki, P. & Essex, M. in *AIDS: Clinical and Research Perspectives* 63-73 (Dekker, New York, 1987).
- Ohta, Y. *et al. Int. J. Cancer* (in the press).
- Kanki, P. *et al. Science* **232**, 238-243 (1986).
- Clavel, F. *et al. Science* **233**, 343-346 (1986).
- Albert, J. *et al. Aids Res. and Human Retroviruses* **3**, 3-10 (1987).
- Kanki, P. *et al. Science* **236**, 827-831 (1987).

An H⁺-ATPase in opposite plasma membrane domains in kidney epithelial cell subpopulations

Dennis Brown*, Sheldon Hirsch†‡ & Stephen Gluck‡§

* The Renal Unit, Massachusetts General Hospital and Department of Pathology, Harvard Medical School, Boston, Massachusetts 02114, USA

† Department of Medicine, Washington University School of Medicine, St Louis, Missouri, 63130, USA

‡ The Renal Division, Jewish Hospital, St Louis, Missouri, 63110, USA

§ Department of Medicine and Cell Biology and Physiology, Washington University School of Medicine, St Louis, Missouri, 63130, USA

Vectorial solute transport by epithelia requires the polarized insertion of transport proteins into apical or basolateral plasmalemmal domains. In the specialized intercalated cells of the kidney collecting duct, the selective placement of an apical plasma membrane proton-pumping ATPase (H⁺-ATPase) and of a basolateral membrane anion-exchange protein results in transepithelial proton secretion¹. It is currently believed that amino-acid sequences of membrane proteins contain critical signalling regions involved in sorting these proteins to specific membrane domains². Recently, it was proposed that intercalated cells can reverse their direction of proton secretion under different acid-base conditions by redirecting proton pumps from apical to basolateral membranes, and anion exchangers from basolateral to apical membranes³. But others have found that antibodies raised against the red cell anion-exchange protein (Band 3) only labelled intercalated cells at the basolateral plasma membrane^{4,5}, providing evidence against the model of polarity reversal. In this report, we have examined directly the distribution of proton pumps in kidney intercalated cells using specific polyclonal antibodies against subunits of a bovine kidney medullary H⁺-ATPase. We find that some cortical collecting duct

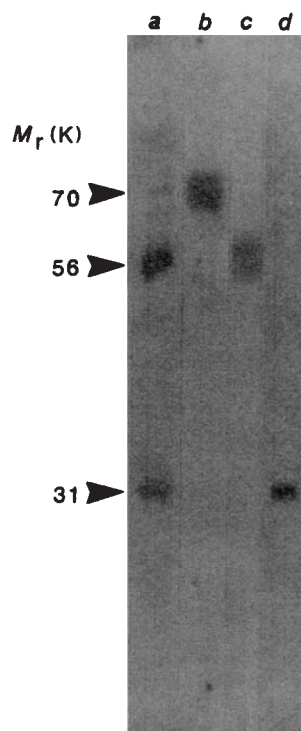


Fig. 1 Immunoblots of purified bovine kidney H⁺-ATPase probed with affinity-purified polyclonal antibodies as indicated: lane *a*, whole anti-holoenzyme antiserum; lane *b*, anti-70K; lane *c*, anti-56K; lane *d*, anti-31K.

Methods. H⁺-ATPase (500 µg) was separated by SDS-PAGE and transferred electrophoretically to nitrocellulose paper²³. Nitrocellulose strips of individual subunits were cut from the sheet. The strips were incubated first with 5% BSA, 0.2% Tween in phosphate buffered saline (PBS) for 1 h at 45 °C to block non-specific binding, and then with whole antiserum diluted 1:10 in the same blocking solution. Strips were washed three times with 0.2% Triton X-100 in PBS, and three times with PBS. Antibodies were eluted with 10 mM glycine-HCl, 10 mM sodium azide, pH 2.5, and neutralized in 1 M Tris. Immunoblots were performed on nitrocellulose strips containing the electrophoretically separated enzyme (5 µg). Strips were incubated in 5% BSA in PBS at 45 °C to block non-specific binding, and then incubated with primary antibody diluted in blocking solution for 2 h. Strips were then washed three times with 0.2% Triton X-100 in PBS, once in PBS containing 1% Triton, 2 M urea and 0.1% glycine, and six times with PBS. Strips were reincubated in blocking solution for 5 min at room temperature, and then with 500,000 c.p.m. ¹²⁵I-protein A in 1 ml blocking solution for 1 h at room temperature. The strips were washed three times with 0.2% Triton in PBS and three times with PBS, and binding was visualized by autoradiography.

intercalated cells have apical plasma membrane proton pumps, whereas others have basolateral pumps. This is the first direct demonstration of neighbouring epithelial cells maintaining opposite polarities of a transport protein. Thus, either subtle structural differences exist between proton pumps located at opposite poles of the cell, or factors other than protein sequence determine the polarity of H⁺-ATPase insertion.

Kidneys from adult male Sprague-Dawley rats anaesthetized with Inactin were fixed by perfusion with paraformaldehyde/lysine/periodate⁶. Some pieces were embedded in LX 112 epoxy resin without osmification for semithin sectioning, and other pieces were embedded in Lowicryl K4M by a rapid-embedding technique⁷. Immunostaining was performed either on semithin sections from which the LX 112 was removed before staining⁸, or on Lowicryl ultrathin sections with no prior etching procedure.

Bovine medullary H⁺-ATPase was purified on a monoclonal-antibody immunoaffinity column as previously described⁹. The

Fig. 2 *a*, Semithin section of collecting duct from the inner stripe of the outer medulla, fixed by perfusion with paraformaldehyde-lysine-periodate. Following removal of the resin⁸, the section was incubated for 2 h with a specific antibody against the 56K subunit of a bovine medullary H^+ -ATPase; after rinsing, sections were incubated for 1 h with goat anti-rabbit IgG coupled to FITC (Calbiochem). All intercalated cells of the collecting duct are heavily labelled at their apical pole, reflecting labelling of the apical plasma membrane and a population of vesicles in the apical cytoplasm. Adjacent principal cells are only weakly stained. Scale bar, 20 μ m. *b*, Semithin section of a cortical collecting duct, incubated as in Fig. 2*a* with a specific antibody against the 56K subunit of the H^+ -ATPase. In addition to cells with an apical labelling (arrows), some cells show a predominant basolateral labelling (double arrows) whereas others have a diffuse labelling that is not restricted to one pole of the cell (arrowheads). Scale bar, 20 μ m. *c*, Thin section of an intercalated cell from kidney cortex embedded in Lowicryl K4M and incubated to show H^+ -ATPase antigenic sites, using the protein A-gold technique. Thin sections were incubated for 2 h on a drop of specific anti-56K antibody, rinsed, and incubated for 45 min with protein A-gold solution prepared from 8-nm gold particles²⁴.

The insets show details of the apical (upper) and basolateral (lower) poles of the same cell. In this cell, the apical membrane is heavily labelled with gold particles, concentrated on the cytoplasmic side of the membrane, whereas the basolateral membrane is virtually unlabelled. Scale bar, 1.0 μ m (insets, 0.1 μ m). *d*, Thin section of cortical intercalated cell, incubated as in Fig. 2*c*, showing a cell with basolateral antigenic sites for H^+ -ATPase. The insets show details of the unlabelled apical (upper) plasma membrane and the heavily labelled basolateral (lower) plasma membrane. Scale bar, 1.0 μ m (insets, 0.1 μ m).

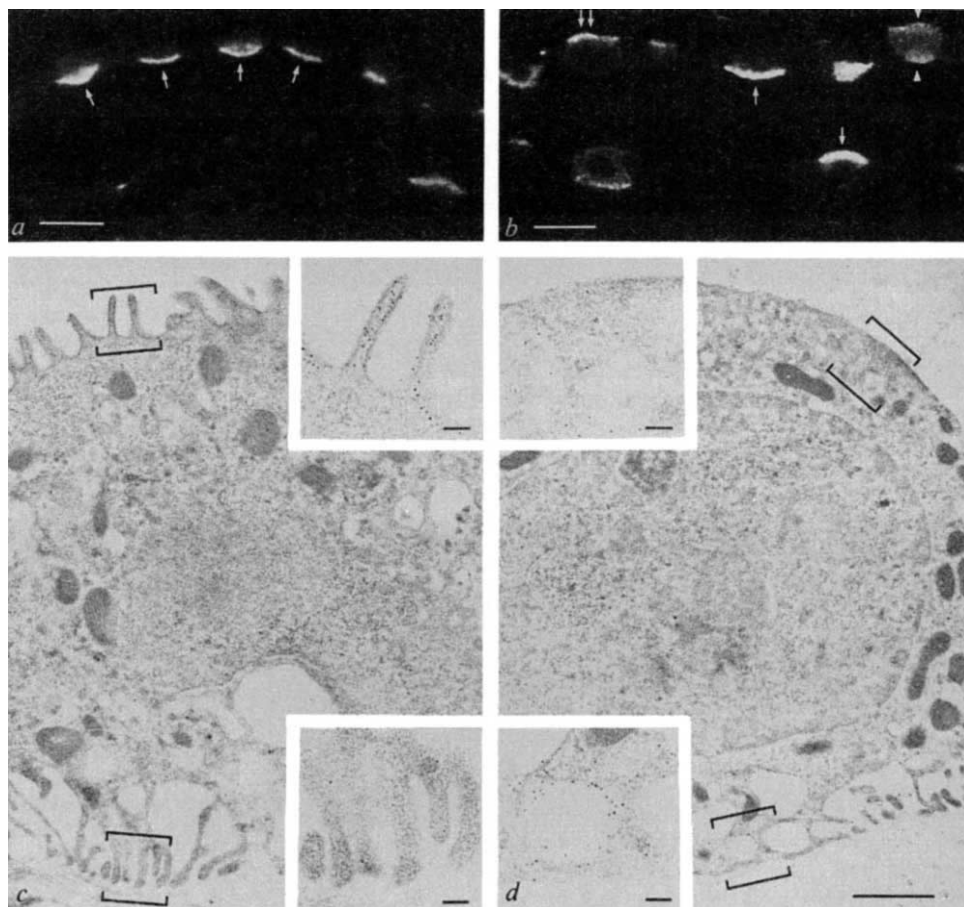


Fig. 3 Thin section of a region of cytoplasm from the middle portion of a diffusely labelled intercalated cell that showed labelled vesicles throughout the cytoplasm, but little or no gold-particle labelling of either the apical or basolateral plasma membrane. Many antigenic sites detected with the anti-56K antibody are associated with cytoplasmic vesicles (arrows), but some labelling appears unrelated to vesicles. This could represent free cytoplasmic subunits of the enzyme or reflect surface labelling of tangentially sectioned vesicles. Scale bar, 0.25 μ m.



purified protein is a multisubunit enzyme which, when reconstituted into liposomes, shows ATP-dependent acidification⁹. Subunits of relative molecular mass (M_r) 31,000 (31K), 56K and 70K comprise about 75% of the enzyme mass by gel densitometry. These subunits are extrinsic membrane proteins that are removed from intact kidney vesicles by washing with buffer at pH 11.3 (not shown), and they form part of the cytoplasmic domain of the H^+ -ATPase (ref. 10). Polyclonal antisera to the holoenzyme were raised in New Zealand white rabbits, and specific antibodies to the 31, 56 or 70K subunits were purified by immunoabsorbing the anti-holoenzyme serum with nitrocellulose strips containing specific H^+ -ATPase subunits separated electrophoretically. The specificity of the affinity-purified antibodies is shown in Fig. 1.

Medullary intercalated cells in the inner stripe showed strong apical labelling by immunofluorescence with all three antibodies (Fig. 2*a*). In the cortex, however, many cells were labelled at their basal pole (Fig. 2*b*), and others had a diffuse, non-polarized

labelling (Fig. 2*b*). All of these cells were identified as intercalated cells by electron microscopy. The protein A-gold technique¹¹ showed antigenic sites concentrated on the cytoplasmic side of either the apical or basolateral plasma membranes (Fig. 2*c*, *d*). Cells with a heavy apical labelling showed little or no basolateral labelling, and vice versa. In both cases the label was seen overlying a dense membrane-coating material that is unrelated to clathrin¹², but is morphologically identical to the cytoplasmic domain of purified bovine kidney H^+ -ATPase (ref. 10). In the cells with a diffuse-staining pattern, label was associated primarily with vesicles distributed throughout the cytoplasm (Fig. 3), although some gold-particle labelling of the cytoplasm could not be excluded. These vesicles participate in endo- and exocytosis, and are believed to carry proton pumps to and from the cell surface¹³⁻¹⁷.

The percentages of antibody-labelled cells in different regions of the collecting duct were: cortex, 41% ($n = 2,111$); outer stripe of outer medulla, 37% ($n = 2,159$); inner stripe of outer medulla,

Table 1 Percentage of intercalated cells in cortical collecting duct showing apical, basolateral or diffuse staining for H⁺-ATPase

Animal	Antibody	Staining pattern		
		Apical	Basolateral	Diffuse
1 (86)	70K	31	9	60
2 (151)	70K	40	12	48
3 (77)	70K	44	38	18
4 (131)	56K	42	13	45
5 (127)	56K	39	7	54
6 (89)	56K	30	52	18
7 (150)	31K	49	15	36
8 (63)	31K	35	54	11
9 (78)	31K	50	6	44

Numbers in parentheses indicate the number of cells counted to derive each percentage. Cells were counted on sections immunostained by the ABC-peroxidase procedure (Vector, Burlingame, California).

42% ($n = 3,927$); inner medulla (initial part), 13% ($n = 1,965$). These figures are consistent with previous estimations of the prevalence of intercalated cells in each collecting-duct segment; the values were similar with all three antibodies. In addition, we quantified the percentage of intercalated cells with the three distinct labelling patterns in cortical collecting ducts from different animals (Table 1). The proportion of cells with basolateral and diffuse staining appeared to be inversely related, and varied considerably among animals, whereas the percentage of cells with marked apical staining was less variable. These findings suggest that redistribution of proton pumps from a diffuse to a basolateral location may occur, whereas the population of cells with apical pumps is more stable. They do not exclude, however, a relocation of proton pumps from basolateral to apical plasma membranes, or vice versa.

Our results provide the first direct evidence in support of models of bicarbonate secretion by the cortical collecting tubule, in which proton pumps located basolaterally are believed to generate intracellular bicarbonate that exits at the apical plasma membrane^{1,3,18}. The heavy apical staining of all intercalated cells in the inner stripe of the outer medulla is consistent with physiological data showing that this segment has the highest rate of transepithelial proton secretion in the kidney¹⁹.

What are the mechanisms that enable the cortical intercalated cells to direct proton pumps to either the apical or basolateral plasma membrane, or to intracellular vesicular compartments? One possibility is that proton pumps destined for the different membrane domains have alternative structures that provide targeting information. Our studies demonstrate that apical and basolateral pumps are immunologically similar, but more investigation will be necessary to determine whether subtle structural differences are present among pumps in different membrane compartments. Indeed, preliminary evidence does suggest that both the 56K (ref. 9) and 31K subunits may have isoforms (S.H. *et al.*, unpublished data). Alternatively, factors other than enzyme structure may determine cellular destination. It has been suggested^{20,21} that the internal pH of transport vesicles can determine their polarity of fusion, but both apical and basolaterally directed vesicles appear to be highly acidic in intercalated cells³. The microtubular system could also be involved in sorting processes. A viral haemagglutinin is mis-directed after colchicine treatment of MDCK cells²², and microtubule disruption prevents the rapid, CO₂-induced insertion of H⁺-ATPase into the apical membrane of proton-secreting cells^{16,17}; whether microtubules can regulate the polarity of proton pump insertion remains to be investigated. In conclusion, our data provide a unique example of adjacent epithelial cells with opposing polarities with respect to a major transporting enzyme. These cells represent an intriguing system in which factors that are involved in the specific targeting and insertion of membrane proteins can be examined.

This project was supported by the NIH, the National Kidney Foundation (S.H.) and the Searle Scholars' Program (S.G.). D.B. is an Established Investigator of the American Heart Association and received partial support from a Chugai Fellowship and a grant from the Milton Fund. We thank Betty Ytzaina for secretarial help and Joanne Natale for excellent technical assistance.

Received 30 November 1987; accepted 12 January 1988.

- Steinmetz, P. R. *Am. J. Physiol.* **251**, F173-F187 (1986).
- Kelly, R. B. *Nature* **326**, 14-15 (1987).
- Schwartz, G. J., Barasch, J. & Al-Awqati, Q. *Nature* **318**, 368-371 (1985).
- Drenckhahn, D., Schluter, K., Allen, D. P. & Bennett, V. *Science* **230**, 1287-1289 (1985).
- Schuster, V. L., Bonsib, S. M. & Jennings, M. L. *Am. J. Physiol.* **251**, C347-C355 (1986).
- McLean, I. W. & Nakane, P. K. *J. Histochem. Cytochem.* **22**, 1077-1083 (1974).
- Altman, L. G., Schneider, B. G. & Papermaster, D. S. *J. Histochem. Cytochem.* **32**, 1217-1223 (1984).
- Maxwell, M. H. *J. Microsc.* **112**, 253-255 (1978).
- Gluck, S. & Caldwell, J. *J. biol. Chem.* (in the press).
- Brown, D., Gluck, S. & Hartwig, J. H. *J. Cell Biol.* **105**, 1637-1648 (1987).
- Roth, J., Bendayan, M. & Orci, L. *J. Histochem. Cytochem.* **26**, 1074-1081 (1978).
- Brown, D. & Orci, L. *Am. J. Physiol.* **250**, C605-C608 (1986).
- Madsen, K. M. & Tisher, C. C. *Am. J. Physiol.* **245**, F670-F679 (1983).
- Brown, D., Weyer, P. & Orci, L. *Anat. Rec.* **218**, 237-242 (1987).
- Schwartz, G. J. & Al-Awqati, Q. *J. clin. Invest.* **75**, 1638-1644 (1985).
- Stetson, D. L. & Steinmetz, P. R. *Am. J. Physiol.* **245**, C113-C120 (1983).
- Gluck, S., Cannon, C. & Al-Awqati, Q. *Proc. natl. Acad. Sci. U.S.A.* **79**, 4327-4331 (1982).
- Stetson, D. L. & Steinmetz, P. R. *Am. J. Physiol.* **249**, F553-F565 (1985).
- Lombard, W. E., Kokko, J. P. & Jacobson, H. R. *Am. J. Physiol.* **244**, F289-F296 (1983).
- Orci, L., Ravazzola, M. & Anderson, R. G. W. *Nature* **326**, 77-79 (1987).
- Caplan, M. J. *et al. Nature* **329**, 632-635 (1987).
- Rindler, M. J., Ivanov, I. E. & Sabatini, D. D. *J. Cell Biol.* **104**, 231-241 (1987).
- Towbin, M., Staehelin, T. & Gordon, J. *Proc. natl. Acad. Sci. U.S.A.* **76**, 4350-4354 (1979).
- Slot, J. W. & Geuze, H. J. *Eur. J. Cell Biol.* **38**, 87-93 (1985).

ICAM, an adhesion ligand of LFA-1, is homologous to the neural cell adhesion molecule NCAM

David Simmons, Malegapuru W. Makgoba* & Brian Seed

Department of Molecular Biology and Department of Genetics, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts, 02114, USA

* Immunology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 02115, USA

Antigen-specific cell contacts in the immune system are strengthened by antigen-nonspecific interactions, mediated in part by lymphocyte-function associated (LFA) antigens^{1,2}. The LFA-1 antigen is widely expressed on cells of haematopoietic origin and is a major receptor of T cells, B cells and granulocytes³. LFA-1 mediates the leukocyte adhesion reactions underlying cytolytic conjugate formation, helper T-cell interactions, and antibody-dependent killing by natural killer cells and granulocytes. Recently, ICAM-1 (intercellular adhesion molecule-1) has been defined as a ligand for LFA-1⁴⁻⁶. Monoclonal antibodies to ICAM-1 block T lymphocyte adhesion to fibroblasts and endothelial cells and disrupt the interaction between cytotoxic T cells and target cells. In addition, purified ICAM-1 reconstituted into artificial membranes binds LFA-1⁺ cells⁶. ICAM-1 is found on leukocytes, fibroblasts, epithelial cells and endothelial cells and its expression is regulated by inflammatory cytokines. LFA-1 has been placed in the integrin family of cell surface receptors by virtue of the high sequence similarity between the LFA-1 and integrin β chains^{7,8}. The adhesion ligands of the integrin family are glycoproteins bearing the Arg-Gly-Asp (RGD) sequence motif, for example, fibronectin, fibrinogen, vitronectin and von Willebrand factor⁹. Here we show that a complementary DNA clone ICAM-1 contains no RGD motifs, but instead is homologous to the neural cell adhesion molecule NCAM^{10,11}.

A cDNA library was constructed using RNA prepared from HL-60 cells induced with 12-*O*-tetradecanoyl phorbol 13-acetate

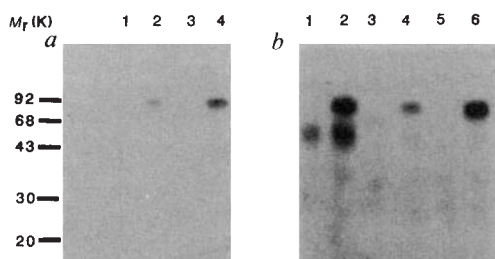


Fig. 1 Immunoprecipitation of ICAM-1 from ^{125}I -surface-labelled cells, using mAb 84H10. *a*, Mock-transfected (lane 1) and pICAM-1-transfected (lane 2) COS cells. Uninduced (lane 3), and TPA-stimulated (lane 4) HL-60 cells. *b*, Regulation of ICAM-1 synthesis in HL-60 cells by cytokines. Lane 1, γ -IFN-induced, without antibody; lane 2, γ -IFN-induced, with antibody; lane 3, TNF-induced, without antibody; lane 4, TNF induced, with antibody; lane 5, IL-1 β induced, without antibody; lane 6, IL-1 β induced, with antibody. COS cells were transfected with pICAM-1 as previously described¹². HL-60 cells were maintained at $5 \times 10^5 \text{ ml}^{-1}$ and test reagents added for 48 h at the following concentrations; 50 ng ml^{-1} TPA, 100 U ml^{-1} γ -IFN, 200 U ml^{-1} TNF and 10 U ml^{-1} IL-1 β .

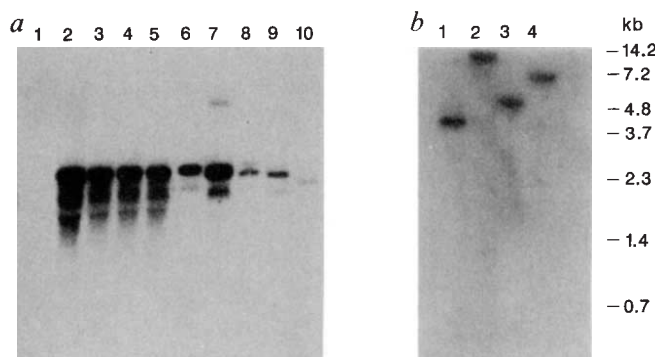


Fig. 2 Expression and structure of ICAM-1 gene. *a*, RNA blot hybridization. Total RNA (20 μg) was denatured in formaldehyde, electrophoresed, transferred to nylon membranes and hybridized with ICAM-1 cDNA. Lane 1, uninduced HL-60; lane 2, TPA-induced HL-60; lane 3, γ -IFN-induced HL-60; lane 4, IL-1 β -induced HL-60; lane 5, TNF-induced HL-60; lane 6, JY (lymphoblastoid cell line); lane 7, Raji (Burkitt's lymphoma); lane 8, Peer (T cell leukaemia); lane 9, T blasts; lane 10, LAK cells. The 1.9 kb species is partially occluded by the 18S rRNA. Stimulation of HL-60 cells was the same as in Fig. 1. T-cell blasts were obtained by culture of peripheral blood lymphocytes with phytohaemagglutinin for 48 h. *b*, Genomic DNA blot hybridization. Human placenta genomic DNA (20 μg) was digested with restriction enzymes, electrophoresed, transferred to nylon membranes and hybridized with ICAM cDNA. Lane 1, *Eco*RI; lane 2, *Bam*HI; lane 3, *Hind*III; lane 4, *Dra*I.

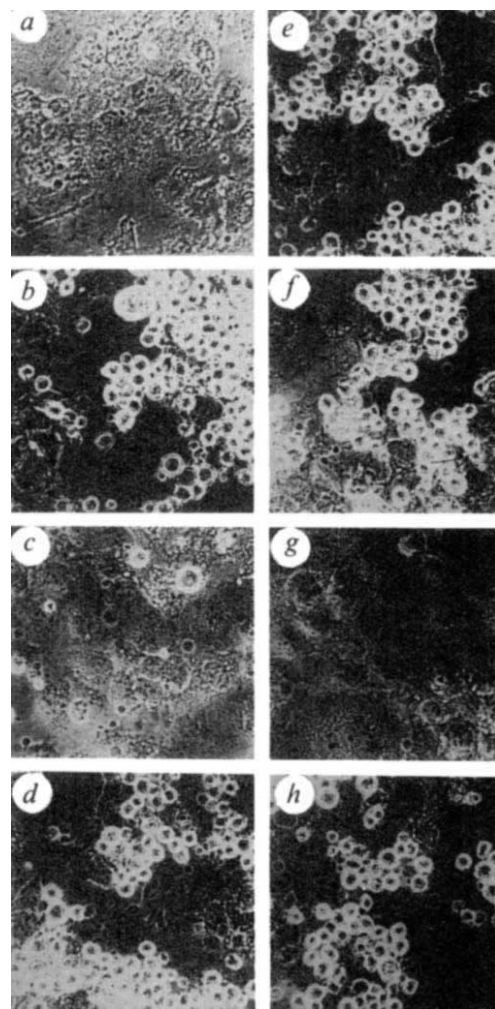


Fig. 3 Functional analysis of pICAM-1. COS cells were transfected with control vector πH3M or pICAM-1 cDNA and HL-60 cell adhesion assays performed 48 h later. *a*, COS cells transfected with vector only; *b-h*, COS cells transfected with pICAM-1 and treated as indicated. *b*, no antibodies present; *c*, anti-ICAM antibody on COS cells; *d*, control anti-HLA antibody on COS cells; *e*, anti-ICAM antibody on HL-60 cells; *f*, anti-HLA antibody on HL-60 cells; *g*, anti-LFA-1 antibodies on HL-60 cells; *h*, anti-LFA-1 antibodies on COS cells. COS cells or HL-60 cells were preincubated with the stated antibodies at 1 $\mu\text{g ml}^{-1}$ for 1 h at 4 $^{\circ}\text{C}$ in PBS/5% fetal bovine serum. In *c*, the anti-ICAM-1 antibody (84H10) was washed off the COS cells to prevent the excess from binding the HL-60 cells. In all other cases the antibodies were left in contact with the cells to prevent leaching. The HL-60 cells were allowed to adhere to the COS cells for 30 min at 37 $^{\circ}\text{C}$ in medium containing 10 mM Mg^{2+} , washed 3 times and photographed.

(TPA). The library was transfected into COS cells and cells expressing surface antigens were recovered by panning with the anti-ICAM monoclonal antibodies (mAbs) 8F5 and 84H10 (ref. 12). Episomal DNA was recovered from the panned cells and the expression-panning cycle^{13,14} repeated twice to obtain a cDNA clone designated pICAM-1.

COS cells transfected with pICAM-1 gave positive surface immunofluorescence reactions with three anti-ICAM-1 antibodies; 8F5 and 84H10 (data not shown), and RR-1 (N. Hogg, personal communication). Immunoprecipitation of pICAM-1 transfected COS cells with 84H10 antibody gave a band of relative molecular mass (M_r) 100,000 (100K, Fig. 1). A slightly larger protein (110K) was precipitated from HL-60 cells induced for 48 h with either TPA, γ -interferon (γ -IFN), tumour necrosis

factor (TNF), or interleukin-1 β (IL-1 β), but was absent from uninduced cells (Fig. 1). The smaller molecular mass of ICAM-1 expressed in COS cells is consistent with the reduced molecular masses observed for other surface antigens expressed in COS cells^{13,14}.

RNA blot analysis (Fig. 2a) showed two species of 3.2 kilobases (kb) and 1.9 kb present in HL-60 cells stimulated with either TPA, γ -IFN, TNF or IL-1 β , but absent in uninduced cells. The 1.9-kb band is occluded by the 18S ribosomal RNA but is present as a distinct species in poly(A)⁺ RNA (data not shown). Thus, the expression of ICAM-1 is regulated by a number of inflammatory cytokines, apparently at the transcription level. Similar species were present in B cells (JY and Raji), and T cells (Peer and T blasts). A slightly smaller transcript was

receptor interaction⁹. But ICAM-1 contains no RGD motifs, bearing instead a single RGE sequence at position 152. A search of the National Biomedical Research Foundation¹⁶ (NBRF) database revealed no significant similarities to other proteins. Comparison with a laboratory database containing recently published surface proteins, however, did reveal a surprising and significant similarity between ICAM-1 and the neural cell adhesion molecule NCAM-1 (refs 10, 11; Fig. 4b). The optimal alignment score obtained using the NBRF ALIGN program is eight standard deviations above the mean score obtained from 500 random permutations of the sequences. The probability of the spontaneous occurrence of an equal or higher score is $\sim 10^{-9}$. Using a database of known immunoglobulin-related sequences it has been shown that ICAM-1 may be divided into five domains (28–112, 115–206, 217–310, 312–391, and 399–477), each of which shows significant similarity with other members of the immunoglobulin superfamily¹⁷ (A. F. Williams, personal communication). For example, domain I is similar to CD3, whereas domains IV and V are similar to domains of myelin-associated glycoprotein¹⁸ and carcinoembryonic antigen¹⁹. All five domains of NCAM align with the domains in ICAM-1 (Fig. 4b), and the principal contribution to the similarity comes from domains II and III of ICAM-1. Finally, the T-cell-specific adhesion molecule CD2 shows roughly the same similarity to NCAM as does ICAM, but ICAM and CD2 are only weakly related (not shown). Thus, some precursor of NCAM is ancestral to both ICAM and CD2.

The similarity of ICAM and NCAM is particularly interesting as it brings together lymphoid and neuronal adhesion molecules. In addition, it is surprising that ICAM is immunoglobulin-related, as its receptor LFA-1 is not. The LFA-1/ICAM-1 pairing demonstrates the interaction of two distinct molecular families, the immunoglobulin family and the integrin family. The availability of a functional ICAM-1 cDNA will allow a better assessment of the importance of ICAM-1/LFA-1-mediated adhesion in antigen-specific leukocyte function, including T-cell mediated killing, T-helper responses and antibody-dependent cell-mediated killing.

We thank Andrew J. McMichael for the myeloid panel of antibodies from the Third International Workshop on Leukocyte Typing, Alan F. Williams for detailed comparison of ICAM with the immunoglobulin superfamily, Nancy Hogg for cytometric analysis of ICAM-1 transfected COS cells and information about mAbs 8F5 and 84H10, Michael Bevilacqua for suggesting the use of HL-60 cells in the adhesion assays and for IL-1 and TNF, Ivan Stamenkovic for B-cell RNAs, and Mike Cherry and Jennifer Doudna for oligonucleotides. This work was supported by a grant from Hoechst AG.

Received 15 October; accepted 22 December 1987.

1. Springer, T. A., Dustin, M. L., Kishimoto, T. K. & Marlin, S. D. *A. Rev. Immun.* **5**, 223–252 (1987).
2. Anderson, D. C. & Springer, T. A. *Rev. Med.* **5**, 175–194 (1987).
3. Rothlein, R. & Springer, T. A. *J. exp. Med.* **163**, 1132–1149 (1987).
4. Rothlein, R., Dustin, M. L., Marlin, S. D. & Springer, T. A. *J. Immun.* **137**, 1270–1275 (1986).
5. Dustin, M. L., Rothlein, R., Bahn, A. K., Dinarello, C. A. & Springer, T. A. *J. Immun.* **137**, 245–254 (1986).
6. Marlin, S. D. & Springer, T. A. *Cell* **51**, 813–819 (1987).
7. Kishimoto, T. K., O'Connor, K., Lee, A., Roberts, T. M. & Springer, T. A. *Cell* **48**, 681–690 (1987).
8. Hynes, R. O. *Cell* **48**, 549–554 (1987).
9. Ruoslahti, E. & Pierschbacher, M. D. *Cell* **44**, 517–518 (1987).
10. Cunningham, B. A. *et al. Science* **236**, 799–806 (1987).
11. Barthels, D. *et al. EMBO J.* **6**, 907–914 (1987).
12. *Leukocyte Typing III. White Cell Differentiation Antigens* (eds McMichael, A. J. *et al.*) (Oxford University Press, 1987).
13. Seed, B. & Aruffo, A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3365–3369 (1987).
14. Aruffo, A. & Seed, B. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8573–8577 (1987).
15. von Heijne, G. *Nucleic Acids Res.* **14**, 4683–4690 (1986).
16. Dayhoff, M. O., Barker, W. C. & Hunt, L. T. *Meth. Enzym.* **91**, 524–545 (1983).
17. Williams, A. F. *Immun. Today* **8**, 298–303 (1987).
18. Arguim, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 600–604 (1987).
19. Beauchemin, N., Benchimol, S., Cournoyer, D., Fuks, A. & Stanners, C. P. *Molec. cell. Biol.* **7**, 3221–3230 (1987).
20. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).

The adult T-cell receptor δ -chain is diverse and distinct from that of fetal thymocytes

John Francis Elliott*, Edwin P. Rock*,
Phillip A. Patten†, Mark M. Davis*†
& Yueh-hsiu Chien*

* Departments of Medical Microbiology and † Biology,
and † Howard Hughes Medical Institute, Stanford University,
Stanford, California 94305, USA

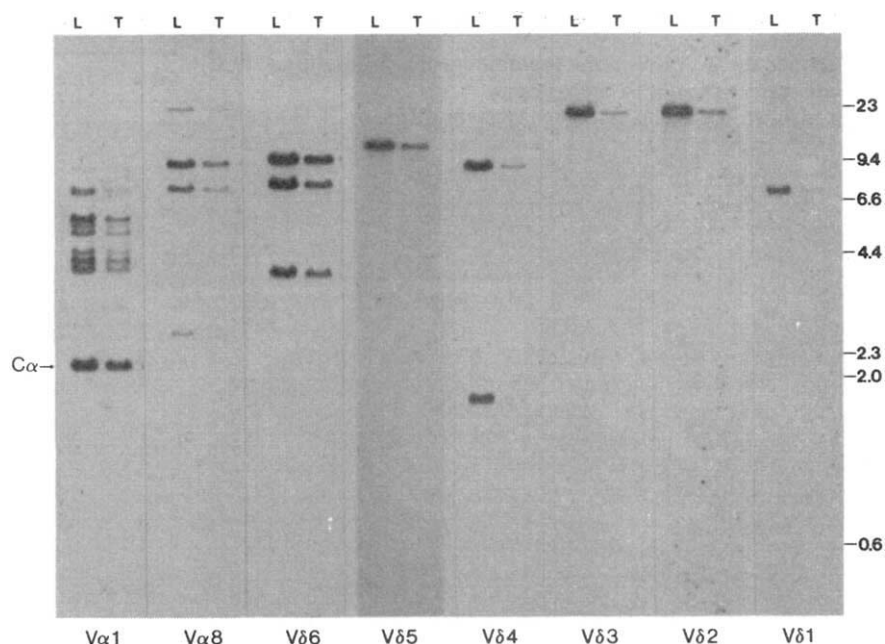
T lymphocytes recognize foreign molecules using the T-cell receptor (TCR), a disulphide-linked heterodimer closely associated with the CD3 polypeptide complex on the cell surface. The TCR $\alpha\beta$ heterodimers seem largely responsible for the recognition properties of both helper (T_H) and cytotoxic (T_C) T cells^{1,2}. Recently, a second CD3-associated T-cell receptor heterodimer, $\gamma\delta$, has been described^{3–9}. Cells bearing the $\gamma\delta$ receptor appear before those bearing $\alpha\beta$ during thymic ontogeny^{8,9} and persist as a minor component (1–10%) of mature peripheral T cells. Their function is unknown. As there are a limited number of functional TCR V_γ gene segments¹⁰, the size and potential diversity of the V_δ repertoire is important for the number of different antigens that may be recognized by $\gamma\delta$ heterodimers. The δ -chain locus is located 75 kilobases (kb) 5' to the TCR C_α coding region^{11,12}, raising the possibility that the α and δ V-region repertoires may overlap. Also, analysis of rearrangements at the δ -chain locus in developing thymocytes shows distinct fetal and adult patterns¹¹ indicating that there may be differences between the fetal and adult V_δ repertoires. To address these questions, we have characterized a large number of δ -containing complementary DNA clones from adult double-negative thymocytes ($CD4^-8^-$), an immature population that is enriched for $\gamma\delta$ -bearing cells⁵. We find that a limited number of V_δ sequences are used, showing little overlap with known adult V_α s and differing significantly from fetal V_δ s. But as two D elements may participate simultaneously in V_δ gene assembly, and random nucleotides may be added at any one of three junctional points, the potential number of different δ chains that can be made in the adult thymus is very large ($\sim 10^{13}$).

Complementary DNA libraries were constructed using RNA from adult double-negative ($CD4^-8^-$) thymocytes (DN thymocytes). Of the 129 δ -cDNA clones, 53 contained the first domain of C_δ , and 17 of these were sequenced. Ten contain all or part of a V region (Fig. 1a), and these are divided into leader (L), variable (V), diversity (D), and joining (J) regions by comparison with the published V_δ , V_α , $D_{\delta 1}$, $D_{\delta 2}$, $J_{\delta 1}$ and $J_{\delta 2}$ gene segments. Eight of the ten cDNA clones represent potentially functional messages, with the V, D and J elements joined in-frame. This may reflect the fact that a significant fraction of adult double negative cells express $\gamma\delta$ receptors on their surface (5–10%)⁹.

Four of the V sequences are identical (Z44, Z72, Z35 and Z14), indicating that this V_δ gene segment is expressed preferentially in adult double-negative thymocytes. V_δ sequences of DN4 (ref. 11) and Z10 are also identical. The V_δ nucleotide sequence of Z53 is similar (96%) to the Z49 sequence and the previously published V_δ sequence p12 (94%)¹¹; the differences may reflect strain polymorphism (Fig. 1b legend). The V_δ sequence of clone Z80 is identical to that of Z49 except for a large central deletion which spans 134 nucleotides and results in a translational frame shift. Perhaps the Messenger RNA corresponding to Z80 arose from a transcript spliced using signal sequences within the V-region. The V region of Z68, which was found only once in these libraries, is virtually identical (99%) to the previously published V_α TA1 (ref. 13). Z78 (Fig. 1a) contains a VDJ junction which involves both the $D_{\delta 1}$ and $D_{\delta 2}$ gene segments.

Methods. A 900 base-pair (bp) *EcoRI* fragment from the 3' portion of C_{δ}^{11} was ^{32}P -labelled by random hexamer priming²⁹ and used to screen two separate libraries, one in $\lambda\text{gt}10$ (ref. 30), and the other in $\lambda\text{-ZAP}$ (Stratagene). The two libraries were constructed from a single preparation of (unsized) double-stranded cDNA, made by the RNase H method³¹, using as starting material $24\ \mu\text{g}$ of total cellular RNA extracted from BALB/c $\times$ 129 (F_1) adult double-negative thymocytes¹¹. In total $\sim 300,000$ recombinant $\lambda\text{gt}10$ phage and 600,000 recombinant $\lambda\text{-ZAP}$ phage were screened, yielding 129 positive clones. Positive clones were screened secondarily with ^{32}P -labelled DNA probes specific for the 5' region of C_{δ} , as well as probes specific for $J_{\delta}1$ and $J_{\delta}2$ in order to isolate longer and full length clones. Two of the $\lambda\text{gt}10$ clones (designated DN) and 38 of the $\lambda\text{-ZAP}$ clones (designated Z) were subcloned into the plasmid vector bluescript SK minus (Stratagene). Clones which (by mapping) contained cDNA inserts long enough to include V region elements were sequenced using the specific primer-directed dideoxynucleotide chain termination method³².

Fig. 2 Southern blot analysis of mouse liver and thymus DNA, using probes from each of the V_δ subfamilies. Aliquots of 8 μ g of *Eco*RI digested genomic DNA from adult liver (L) or thymus (T) were electrophoresed in a 0.8% agarose gel, blotted onto nitrocellulose and probed with separate V_α or V_δ DNA fragments which had been radiolabelled by random hexamer priming. Blots were hybridized in 1 M NaCl, 50 mM NaPO_4 , pH 6.5 50% formamide, 100 μ g ml^{-1} salmon sperm DNA, 1 $\times$ Denhardt and 10% dextran sulphate at 40 $^\circ\text{C}$ and washed in both 2 $\times$ SSPE and 0.2 $\times$ SSPE at 55 $^\circ\text{C}$. *Hind*III-digested λ phage DNA was used as a size marker.



The 3' end of the V -like element encodes some of the conserved amino acids found in other V_δ genes, but the rest of the element has no open reading frame, suggesting that it arises from a V_δ pseudogene segment. These ten cDNA sequences derive from only five different V_δ gene segments (prototype Z68, DN4, Z44, Z53 and Z49), but the junctions between V and J are all different, indicating that they are derived from transcripts of independent rearrangement events.

Previous results indicate that there are two distinct D elements and two J elements in the mouse δ -chain locus^{11,12}. $J_{\delta 1}$ is found in all ten of the clones described here containing V regions, in six of the seven non-coding transcripts (Fig. 1c) and in 36 of 53 cDNA clones containing the first domain of C_δ , compared to only five clones which use $J_{\delta 2}$. This preferential use of $J_{\delta 1}$ is similar to that observed in fetal thymocyte DNA and in hybridomas derived from fetal cells¹².

Contributions from both D elements at the VDJ junction are evident in most of the adult δ -chain cDNA sequences (Fig. 1a), and diversity of the δ -chain is greatly increased by the apparently random insertion (referred to as 'N-region' addition¹⁴) of as many as six nucleotides between the different gene segments (Fig. 1a). In contrast, the sequences between the V_δ and J_δ segments in fetal thymocytes can be accounted for with $D_{\delta 2}$ alone (Fig. 1a) and there is no discernible $D_{\delta 1}$ or N region contribution. N region addition correlates well with the presence of terminal deoxynucleotidyl transferase (TdT)^{14,15}, and the presence of N regions in adult sequences but not in fetal is

consistent with the low TdT activity found in fetal thymocytes compared to adult cells^{16,17}.

Amino-acid sequences of the five V segments isolated in this study, and the four mouse V_δ sequences reported previously are presented in Fig. 1b, and fall into six subfamilies (subfamily members have >75% sequence homology at the nucleotide level). As with TCR α and β -chain V regions, V_δ sequences are generally 20–60% similar at the amino-acid level (Table 1). We denote these subfamilies $V_{\delta 1}$ – $V_{\delta 6}$ according to their relative deletion frequency in the adult thymus (see below). Z68 is virtually identical to the V_α gene TA1 and V_{M23} shows 94% nucleotide similarity with the V_α gene segment TA27 (ref. 13), but none of the other V_δ gene segments have been found in association with J_α and C_α .

Six of the cDNA clones sequenced do not contain V region sequences. Z17 contains the second J region, $J_{\delta 2}$, spliced to C_δ . Four of the other cDNA clones involve aberrant rearrangements between D_δ and J_δ gene segments (Fig. 1c). DN14, Z76 and Z64 all have a region immediately 5' of $D_{\delta 1}$ joined to $J_{\delta 1}$. The 2–5 extra nucleotides are presumably due to 'N-region' addition. These three sequences differ from the DJC transcripts found in the immunoglobulin heavy chain and TCR β -locus in that none of them contain any of the $D_{\delta 1}$ sequence. Another clone, Z61, joins a gene segment 1.6 kb 3' of $D_{\delta 1}$ to the $D_{\delta 2}$ element and $D_{\delta 2}$ to $J_{\delta 1}$. Here, sequences similar to the consensus heptamer and nonamer found 3' to V and D gene segments are discernible at the 3' side of the putative rearrangement point. Z40 is the only cDNA clone which, like the DJC transcripts of the immunoglobulin heavy-chain and TCR β -chain loci, initiates 5' to $D_{\delta 2}$ and may correspond to a $D_{\delta 2}$ – $J_{\delta 1}$ rearrangement. But the three nucleotides at the 5' side of $D_{\delta 2}$ differ from the germline sequence. This sequence may result from a specific cleavage at the junction of the heptamer and 5' $D_{\delta 2}$ sequences, followed by nuclease digestion, random nucleotide insertion and ligation. As shown in Fig. 1c, none of the sequences have retained a conserved heptamer sequence at their 5' end. Thus it is unlikely that the genomic DNA sequences corresponding to these cDNA clones can be substrates for further site-specific recombination. In contrast, in the immunoglobulin heavy chain and TCR β -chain, DJ joints are thought to be the precursors of VDJ joints^{18–20}. Although recombinations involving the joining of non- V , non- D segments to J have been found in the immunoglobulin heavy-chain locus²¹, they are considerably less frequent than *bona fide* DJ rearrangements. As $J_\delta C_\delta$ is located 5' to $J_\alpha C_\alpha$, the 'aberrant' rearrangement events described here

Table 1 Similarity of V_δ sequences

	V_{M21}	V_{M11}	Z68	DN4	Z44	Z53	Z49	pA12	V_{M23}
V_{M21}	—	24	20	31	28	23	24	22	21
V_{M11}	46	—	24	31	28	35	37	35	31
Z68	47	47	—	28	30	30	32	33	32
DN4	49	52	48	—	34	48	48	49	45
Z44	46	48	49	51	—	34	33	34	32
Z53	46	48	46	61	48	—	92	94	62
Z49	44	48	44	62	47	96	—	87	63
pA12	44	47	44	62	46	97	93	—	66
V_{M23}	44	47	45	62	47	79	78	80	—

The numbers below the diagonal indicate percentage similarity of V region nucleotide sequences; those above the diagonal compare protein sequences. Only that portion of each V region sequence which is shown in Fig. 1b was used in this analysis. To determine percent similarity the Microgenie (Beckman) alignment algorithm was used to align V sequences, and the total number of matched residues was divided by the total length (in residues) of sequences being compared.

may render the δ region non-recombinogenic, promoting subsequent rearrangement to the α locus.

Because at least one V gene segment (Z68 or TA1) can rearrange to either the α or δ locus and the $J_\delta C_\delta$ locus is located just 5' to $J_\alpha C_\alpha$, it is possible that V_δ gene segments are simply V_α s that are closer to the $J_\delta C_\delta$ and $J_\alpha C_\alpha$ loci and are thus more likely to rearrange early in ontogeny into the δ locus. Alternatively, V_δ s and V_α s may be dispersed, and their rearrangements determined by other factors. In an attempt to distinguish between these possibilities, we performed hybridization analysis on total thymocyte and liver DNA using V_δ and V_α gene segments as probes (Fig. 2). If most V_δ elements are located 3' to the V_α s, they should be deleted in the process of α -locus rearrangement (assuming that most V_α s rearrange by deletion), reducing the intensity of V_δ germline band(s) in Southern blots of cells that have rearranged their α -chain genes.

Table 2 shows that V_δ sequences V_{M9} ($V_{\delta 1}$) and V_{M11} ($V_{\delta 2}$), which are rearranged as early as day 14–15 in fetal thymocytes¹², retain only 15% of their germline sequences in adult thymocytes. Z68 (V_α sequence TA1), which is located on the same 17.5-kb *EcoRI* fragment as V_{M11} (data not shown) and is frequently rearranged to the δ locus in day-17 fetal hybridomas, shows an equally diminished intensity in adult cells. Yet V_{M23} ($V_{\delta 6}$), isolated from a day 18–19 fetal thymocyte library¹², shows >94% homology with a V_α sequence (TA27) both in the leader and V gene segments and is only slightly deleted from the adult thymocyte. The other two V_α segments tested, $V_{\alpha 1}$ and $V_{\alpha 8}$, demonstrate a range of relative intensities similar to the $V_{\delta 6}$ family. Thus, known V_δ s and V_α s may fall into either frequency extreme in terms of deletion, and there seems to be no pattern to allow us to distinguish between V_δ gene segments and V_α gene segments.

The correlation of deletional frequency with chromosomal location may be complex because some V genes may rearrange by inversion (ref. 12 and data not shown). It is clear, however, that V_δ s and V_α s may be dispersed in this locus, and that factors other than simple chromosomal location are important in rearrangement.

Fetal and adult double-negative cells are relatively immature cells, with significant developmental potential. The set of V_δ genes found overlaps very little with the known V_α sequences²². However, a day-15 fetal hybrid has a $V_{M9}DJ_\delta$ rearrangement on one chromosome and on the other a V_{M9} gene segment rearranged into a region 3' to C_δ where functional J_α s have been found¹², suggesting that the same V gene may rearrange to both δ and α loci. There may be other examples of overlap between the V_δ and V_α repertoires in T-cell populations which have not yet been characterized.

In contrast to V_α gene segments, where ten of eleven subfamilies cross-hybridize with multiple bands on Southern blots²², the majority of V_δ gene segments cross-hybridize with a single band, and at most with only three bands. This distribution is more similar to that of V_β gene segments (in the mouse genome) where most subfamilies contain just one member^{23,24}.

As four of the eleven V_δ segments studied are virtually identical in sequence, we analysed the frequency of V_δ gene usage in these two libraries by cross-hybridization (Table 2). The complete libraries were screened with all the available V_δ gene segments, and we counted those clones which also cross-hybridized with the first domain of the C_δ gene. A total of 21 cDNA clones (including those presented in Fig. 1a) contained V region sequences. None of the V_δ s isolated from a fetal genomic library were present in the two libraries, suggesting that these V regions are rarely expressed.

Because of the non-random use of V_δ genes in a given cell type, it is difficult to estimate the total number of V_δ genes, but it is likely that only subsets of them are active in adult or fetal T cells, with perhaps 10–12 distinct V regions being used preferentially in a given cell type. This is considerably fewer than the 25–30 estimated V_β gene segments or ~75 V_α segments, but

Table 2 Intensity of V_δ -hybridizable bands in adult thymocyte DNA

V subfamily	Probe	Band Size	Relative Intensity	Frequency of Usage
$V_{\delta 1}$	M9	7.2 kb	0.14	0/21
$V_{\delta 2}$	M11	17.5 kb	0.14	0/21
$V_{\delta 3}$	Z68	17.5 kb	0.15	1/21
$V_{\delta 4}$	DN4	8.4 kb	0.23	4/21
		1.8 kb	0.25	
$V_{\delta 5}$	Z44	10.7 kb	0.43	11/21
$V_{\delta 6}$	Z53	9.4 kb	1.0	5/21
		7.4 kb	0.63	
		3.8 kb	0.83	
$V_{\alpha 1}$	—	—	0.66–0.74	
$V_{\alpha 8}$	—	—	0.67–0.87	

The Southern blots in Fig. 2 were analysed by scanning densitometry. Intensity of each band was estimated by determining the area under each densitometric peak. Areas for corresponding bands in thymus and liver were compared directly but standardized for the amount of DNA in each lane via multiplication by a correction factor. The correction factor was obtained by determining the relative intensity of the C_α signal (indicated) and of mouse chondroitin sulphate core protein sequences (giving bands at 10.7, 7.8, 6.0, 2.8 and 2.3 kb, not shown). Band sizes and their relative intensities in thymus versus liver are shown. The complete cDNA libraries were screened with a probe specific for the first domain of C_δ ¹¹. The 21 positive clones identified were screened with DNA probes from the various V_δ gene fragments. Frequency is given as the number of positive clones for a particular V_δ segment out of the 21 clones containing the first domain of C_δ .

more than the seven V_γ elements reported to date. Comparing our study of adult thymocytes with those of V gene assembly in fetal thymocytes¹² indicates a transition in V region use during ontogeny. Whereas $V_{\delta 1}$ is rearranged and expressed frequently in fetal cells (ref. 12 and Y.C., unpublished data), we found no $V_{\delta 1}$ transcripts in adult DN thymocytes. Instead, half of the V -region-containing cDNAs (11/21) involve another V region, $V_{\delta 5}$. The other three V regions found in fetal thymocytes ($V_{\delta 2}$, $V_{\delta 3}$, and a member of $V_{\delta 6}$ (V_{M23})) could be at most a minor constituent. It is interesting that some V_γ sequences also appear only in fetal cells¹⁰, thus both γ and δ V -region use may change between fetal and adult cells. These data support the idea that the fetal and adult thymuses are populated by distinct types of cells that may be subjected to very different influences²⁵, and raise an intriguing question as to the nature of thymic ligands for $\gamma\delta$ -bearing cells at these different stages of development.

In fetal δ -chain genes $VD_{\delta 2}J_\delta$ joints without N region addition are observed exclusively¹², but in adult cells most transcripts have components from both $D_{\delta 1}$ and $D_{\delta 2}$, with extra nucleotides (N regions) between the different gene elements. This indicates that the fetal δ chains are considerably less diverse, and may recognize a more limited range of epitopes than the adult.

Because there are two D_δ regions, as well as potential N -region additions at each of the three junctions (V - $D_{\delta 1}$, $D_{\delta 1}$ - $D_{\delta 2}$, $D_{\delta 2}$ - J_δ), a very large number of different δ chains could be expressed. The diversity of the TCR δ -chain in adult cells can be estimated by multiplying the number of V regions (~6 in adults), the number of different D and N -region contributions, and the number of J s (two)¹². Joining at the 3' end of V_δ can vary by as many as five nucleotides (as seen for $V_{\delta 5}$ in Fig. 1a) and at the 5' end of $J_{\delta 1}$ by six nucleotides. As different subsets of each D region (ranging from two nucleotides to the entire D segment) may be contributed, there could be 5,000 (55×99) possible combinations of nucleotides from $D_{\delta 1}$ and $D_{\delta 2}$. Adding 0–6 N -region nucleotides (presumed here to be random, although there appears to be some bias towards G and C residues, Fig. 1a) at each of the three junctional points could provide $4^6 + 4^5 + 4^4 + 4^3 + 4^2 + 4^1$ or 5,460 different sequence contributions at each position. Some of this diversity is redundant with that of V , D or J , however, particularly the adjacent

nucleotides. This can be approximated by saying that three out of four times the *N*-region nucleotide adjacent to *V*, *D* or *J* will be different from that of the germline sequence (and thus count towards diversity). This reduces the diversity of each *N* region by ~50% making the total *N* region contribution $(5,450/2)^3$ or 2×10^{10} different sequences. Based on these assumptions the overall diversity in the adult TCR δ chain can be calculated as $[6(V \text{ regions}) \times 2(J \text{ regions}) \times 5(3' V_{\delta} \text{ diversity}) \times 6(5' J_{\delta} \text{ diversity}) \times 5,000(D \text{ region contributions}) \times 2 \times 10^{10}(N\text{-region contributions})] = 3 \times 10^{16}$ possible distinct nucleotide sequences. To convert this to functional protein sequences, three corrections are necessary. One is for the probability that the junctional sequence maintains the reading between *V* and *J*, which will be true for one third of the sequences. The second is for termination codons, which should occur in less than 10% of the in-frame joins, since neither of the two *D* regions contain a termination codon in any of the three reading frames. Finally, due to the redundancy in the genetic code, distinct rearrangement events can give rise to nucleotide sequences that code for identical protein sequences. This occurs principally when an *N*-region residue falls at the third base in a codon, the wobble position. There are, for example, 128 different nucleotide sequences allowed by *N*-region wobble which would give the Z80 protein sequence (Fig. 1a). Since there will be on average three redundant codons at each such wobble position, there will be $(3)^{N/3}$ redundancies for any given junctional sequence, when *N* is the total number of *N*-region residues in that junction. Thus for a junction containing 18 *N*-region residues, six will fall at wobble positions, giving a correction of $(1/3)^6 = 0.001$. Based on these assumptions the amino acid diversity of the TCR δ -chain could be $3 \times 10^{16} \times 1/3$ (in-frame joinings) $\times (1/3)^6$ (codon redundancy) = 10^{13} different sequences, many orders of magnitude higher than any other protein of this type, with almost all of this diversity between the end of *V* and the beginning of *J*. Similar calculations for the TCR γ gene of the mouse, incorporating data showing extensive *N*-region diversification^{26,27}, give an estimate of at least 10^4 different *VJ*_γ sequences. Thus there could be 10^{17} possible $\gamma\delta$ heterodimers, with almost all of the diversity in the *VJ* junction of each chain. As this region is equivalent to the CDR3 loop of antibody *V* regions, the concentration of diversity found here in T-cell receptors may have important implications for all T-cell recognition²⁸.

We thank K. Redman for preparation of the manuscript. Supported by grants from the NIH and the Howard Hughes Medical Institute. J.F.E. is a Centennial Fellow of the Medical Research Council of Canada. E.P.R. and P.A.P. are supported by predoctoral fellowships from NIH. M.M.D. is a Scholar of the PEW Foundation. The sequence data in this publication have been submitted to the EMBL/GenBank Libraries under the accession number Y00298.

Received 5 October; accepted 24 December 1987.

- Dembic, Z. *et al. Nature* **320**, 232-238 (1986).
- Saito, T., Weiss, A., Miller, J., Norcross, M. A. & Germain, R. N. *Nature* **325**, 125-130 (1987).
- Brenner, M. B. *et al. Nature* **322**, 145-149 (1986).
- Bank, I. *et al. Nature* **322**, 179-181 (1986).
- Lew, A. M. *et al. Nature* **324**, 1401-1408 (1986).
- Borst, J. *et al. Nature* **325**, 683-688 (1987).
- Moingeon, P. *et al. Nature* **325**, 723-726 (1987).
- Pardoll, D. M. *et al. Nature* **326**, 79-81 (1987).
- Bluestone, J. A., Pardoll, D., Sharrow, S. O. & Fowlkes, B. J. *Nature* **326**, 82-84 (1987).
- Altman, R. D., Doherty, P. J. & Raulet, D. H. *Cell* **45**, 733-742 (1986).
- Chien, Y., Iwashima, M., Kaplan, K. B., Elliott, J. F. & Davis, M. M. *Nature* **327**, 677-682 (1987).
- Chien, Y. *et al. Nature* **330**, 722-727 (1987).
- Arden, B., Klotz, J., Siu, G. & Hood, L. *Nature* **316**, 783-787 (1985).
- Alt, F. W. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4118-4122 (1982).
- Landau, N. R., Schatz, P. G., Rosa, M. & Baltimore, D. *Molec. cell. Biol.* **7**, 3237-3243 (1987).
- Rothenberg, E. & Triglia, D. *J. Immunol.* **130**, 1627-1633 (1983).
- Penit, C., Jotereau, F. & Gelibert, M. *J. Immunol.* **134**, 2149-2154 (1985).
- Yaoita, Y. *et al. Nucleic Acids Res.* **11**, 7303-7316 (1983).
- Alt, F. W. *et al. EMBO J.* **3**, 1209-1219 (1984).
- Born, W., Yagui, J., Palmer, E., Kappler, J. & Marrack, P. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2925-2929 (1985).
- Yancopoulos, G. D. *et al. EMBO J.* **5**, 3259-3266 (1986).
- Kronenberg, M., Siu, G., Hood, L. E. & Shastri, N. A. *Rev. Immunol.* **4**, 529-591 (1986).
- Patten, P. *et al. Nature* **321**, 40-46 (1984).

- Barth, R. *et al. Nature* **316**, 517-523 (1985).
- Jotereau, F., Heuze, F., Salomon-Vie, V. & Gascan, H. *J. Immunol.* **138**, 1026-1030 (1987).
- Quertermous, T. *et al. Nature* **322**, 184-187 (1986).
- Heilig, J. S. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8070-8074 (1987).
- Davis, M. M. & Bjorkman, P. J. (in preparation).
- Feinberg, A. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).
- Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning Techniques: A Practical Approach* (ed. Glover, D. M.) 49-78 (IRL Press, Oxford, 1984).
- Gubler, U. & Hoffman, B. J. *Gene* **25**, 263-269 (1983).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5436-5440 (1977).

Transposon-induced deletions in *unc-22* of *C. elegans* associated with almost normal gene activity

Jane E. Kiff, Donald G. Moerman,
Lawrence A. Schrieffer & Robert H. Waterston*

Department of Genetics, Washington University School of Medicine,
660 South Euclid Avenue, St Louis, Missouri 63110, USA

The *unc-22* gene of *Caenorhabditis elegans* encodes a protein which is a component of the myosin-containing A-band of the worm's striated body-wall muscle¹⁻³. Among 51 revertants of a transposon-induced mutant, we have identified four which retain a barely detectable mutant phenotype. Molecular analysis shows that three of these have in-frame deletions of 1.0, 1.3 and 2.0 kilobases, whereas the fourth partial revertant and two other apparently complete revertants have small insertions. All these rearrangements involve coding sequence and, in the case of the deletions, result in polypeptides that are shorter than the wild-type protein. The region of the gene containing these rearrangements contains 10 copies of a motif recognized in other regions of the gene (our unpublished data). We suggest that one explanation for the minimally mutant phenotype associated with the deletions is that the size and the repeated nature of the *unc-22* protein structure make it relatively tolerant of substitutions or deletions involving one or a small number of repeated motifs. These results could explain why in some human genetic diseases, such as Duchenne's muscular dystrophy, deletions can be associated with only mild forms of the disease⁴.

The *unc-22* locus is one of more than 30 muscle-affecting genes identified in *C. elegans*^{1,5,6}. Loss-of-function alleles of *unc-22* lead to impaired movement, to a disordered myofilament lattice, and to a distinctive twitching of localized areas in individual muscle cells. This 'twitcher' phenotype is recessive for most alleles under standard conditions, but can be elicited in heterozygotes by placing them in a 1% nicotine solution⁷. Wild-type animals rapidly become stiff and immobile whereas *unc-22* homozygotes or heterozygotes twitch violently for many hours.

The *unc-22* gene has recently been cloned² and found to encode a muscle-specific protein of relative molecular mass (*M_r*) greater than 500,000 (500K) that is localized in the myosin-containing A-bands³. This position, along with previous genetic evidence⁸, strongly suggests that the *unc-22* protein is a component of the thick filament. Its function there is unknown but the twitching phenotype suggests that the protein plays a role in the regulation of contraction.

Because imprecise germline excision of transposable elements can be a rich source of new mutations^{9,10,11}, we examined 51 revertants of a transposon-induced *unc-22* mutant, *st192::Tc1*, for new or subtle phenotypes. To obtain the revertants we used an active mutator strain in which the reversion frequency of *st192::Tc1* is 10^{-3} , several orders of magnitude higher than in non-mutator backgrounds ($<10^{-6}$). We evaluated each of the 51 homozygous revertants for movement, muscle structure and

* To whom correspondence should be addressed.

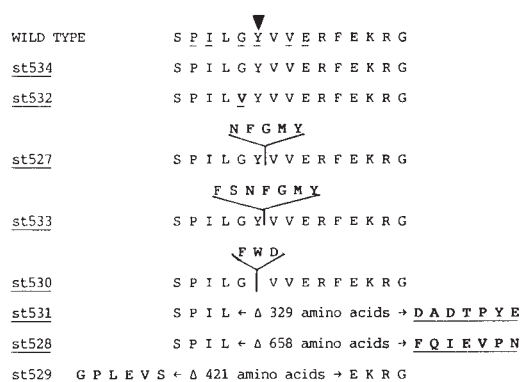


Fig. 1 The amino acid sequence of the wild type in the region surrounding the *st192::Tc1* insertion, and the sequences in the same region from each of the revertants is shown. The original insertion site (arrowhead) is in a tyrosine codon within a conserved region of the repeated motif (see text and Fig. 3). Residues that are strongly conserved in the 14 motifs so far identified are underlined in the wild-type sequence. Residues that are substituted in the revertants are shown in underlined bold type. Inserted residues are also in bold type, above their insertion points. The tyrosine residue absent from *st530* is present in all 14 motifs in the wild-type gene found to date. The deletion mutants (last three lines) disrupt the normal motif by fusing two different motifs out of register (see Fig. 3).

Methods. The wild-type 4.3-kb region was cloned into pUC vectors as a 3.4-kb *Bgl*II fragment and an overlapping 1.0-kb *Sac*I fragment. Smaller restriction fragments were subcloned from these into M13mp18 and M13mp19 and the sequence of both strands was determined by the chain-termination method^{15,16}. The eight revertants and the original insertion mutant, *st192::Tc1*, were cloned into either λ2001 or pUC18/19. The region of interest for each was subcloned into M13mp18/19 and sequenced as above. The amino-acid sequence was determined by conceptual translation of the DNA sequence. Analysis was performed using Microgenie (Beckman) and Staden's DNA-sequence analysis programs¹⁷.

twitching, and all are apparently wild type under normal culture conditions. In 1% nicotine, however, four of the revertants exhibit a weak and transient twitch in the presence of the drug, a phenotype that has not been observed previously for any *unc-22* allele. The phenotypes of three of these partial revertants, *st528*, *st529*, *st531*, are indistinguishable from one another; the fourth, *st530*, exhibits a fainter twitch.

To determine the nature of the molecular changes associated with the reversion events we analysed the mutations in the four partial revertants (*st528*, *st529*, *st530*, *st531*) and in four of the apparently complete revertant alleles (*st527*, *st532*, *st533*, *st534*). Southern analysis showed that three of the four partial revertants have large deletions of 1.0, 1.2 and 2.0 kilobases (kb) in the *unc-22* gene (data not shown). The fourth partial revertant and all four of the others show the wild-type pattern, suggesting precise or nearly precise excision of the 1.6-kb *Tc1* element in these cases.

The *unc-22* gene from each of the revertants was cloned, the sequence was determined in the region of interest and compared to a 4.3-kb region of the wild-type *unc-22* gene spanning the limits of the deletions. Only one revertant, *st534*, restores the wild-type sequence. Another, *st532*, results in a single-residue change, and three (*st530*, *st527* and *st533*) are small insertions of 2–7 amino acids (Fig. 1). As expected, the deletions in *st528*, *st529*, and *st531* maintain the reading frame, but remove 658, 421 and 328 codons respectively (Fig. 1).

We also examined the size of the *unc-22* protein in these eight strains. By immunoblot assay, the protein produced in each of the deficiency strains is detectably smaller than that from the wild type (Fig. 2). Appropriately sized standards for this region of the gel are not available, so that we cannot accurately measure

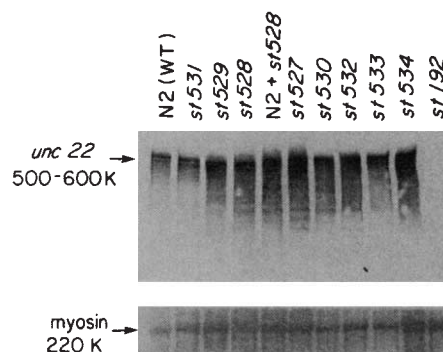


Fig. 2 The upper panel is a Western blot of total protein from wild type (N2) and revertants, reacted with an antibody to the *unc-22* protein. Lane 1, wild-type; lanes 2–4, samples from the mutants with predicted deletions of 328, 421 and 658 amino acids, respectively; lane 5, an equal mixture of extracts from both wild type and the largest-deletion mutant (for direct comparison); lanes 6–10, the other five revertants. The last lane contains a sample from the original *Tc1* mutant, *st192*, which has no detectable full-length *unc-22* product in this assay. The lower panel shows amido-black staining of myosin on the same blot to compare loadings. Each of the deletion revertants contains a detectably smaller protein of comparable abundance to the wild type. The mobility of the *unc-22* protein in the largest-deletion strain is not as high as would be predicted from its size. As the region deleted in this mutant is likely to be entirely coding sequence (see text), the polypeptide could migrate anomalously.

Methods. Animals of the appropriate genotype were collected from standard growth plates and extracts prepared as previously described³. SDS-PAGE was performed using 4% acrylamide¹⁸ then blotted on to nitrocellulose^{3,19}. The filter was treated with the polyclonal antibody R11-3, raised against an *unc-22*-β-galactosidase fusion protein³, then with protein-A-peroxidase. The colour reaction was developed in chloronaphthol and hydrogen peroxide and the transfer subsequently stained with amido-black¹⁹.

the amount of polypeptide deleted. Yet the following evidence strongly suggests the region deleted is entirely coding sequence: the entire 4.3-kb region contains a continuous open-reading frame throughout its length, and the codon usage matches that of other *C. elegans* muscle genes. A search of the region for intron donor and acceptor sites has offered no good candidates. Furthermore, we have shown that in another *Tc1* insertion allele (*st141::Tc1*), in the same site and orientation as *st192::Tc1*, *Tc1* sequences become incorporated into a stable, larger-than-normal messenger RNA. We believe, therefore, that this region consists of one large exon.

The insensitivity of the *unc-22* protein to sequence changes and even to the removal of substantial portions of the protein is surprising, because deletions are usually, although not always, associated with a complete loss of gene activity. Possible explanations for the plasticity of the *unc-22* gene could be either that the region affected is not important to the function of the protein, or that the altered region is largely or totally redundant. Although precedent exists for the first hypothesis¹², our analysis of the *unc-22* sequence from the region around the insertion site leads us to favour the latter suggestion. The sequenced 4.3-kb fragment bears no significant similarity to sequences presently held in data bases, but it does contain 10 copies of a motif (previously recognized by G. M. Benian *et al.*, unpublished results) present in other regions of the gene. The motifs are clustered in groups of two and three throughout the region, as shown in Fig. 3a, and are separated by non-conserved regions. Each motif is slightly less than 100 residues long and contains two conserved domains, designated the A-block and B-block (Fig. 3). The similarity from one repeated motif to another would not be expected if the region were free from selection, and

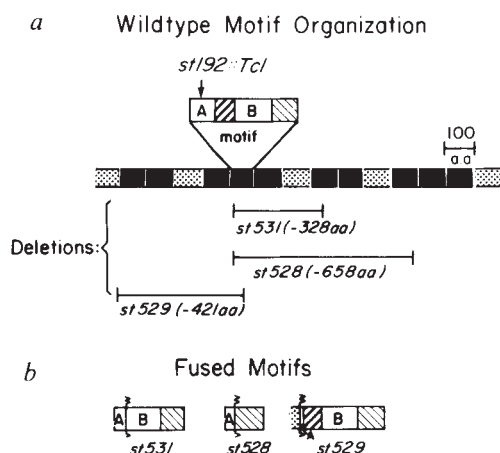


Fig. 3 *a*, The 4.3-kb region spanning the deficiencies contains 10 repeated motifs organized in clusters as shown. For simplicity, the motifs are represented as solid boxes bounded by white vertical lines. A typical motif of just less than 100 amino acids in length is shown in more detail above. The blocks labelled A and B are regions of extensive homology between the 14 copies of the motif so far identified, and the areas with diagonal cross-hatching represent less-conserved regions between the A and B blocks and following the B block. The site of the *st192::Tc1* insertion into this motif is indicated (arrow). The stippled areas between the motif blocks represent the spacer regions, which are not homologous to the motifs but are themselves also about 100 residues in length. The break-points of the deletions *st528*, *st529* and *st531* which accompany *Tc1* excision are shown below. *b*, The aberrant motifs produced by the deletions are illustrated in detail. The deletions produce disruptions of the typical motif, fusing part of one unit to another from either upstream or downstream sites. For example 3 entire repeat units are removed in *st528*, and part of one A block is fused to the less-conserved region to the right of a downstream B block. The amino terminus lies to the left in the diagram.

therefore implies that it has a function. If each of the repeats has a similar function, alterations in any one of them would be buffered by the function of the remaining units. The mild dysfunction associated with the deletions may simply reflect removal of one to three repeats. However, the deletions not only reduce the number of repeats, but also generated an aberrant unit where the deletion end-points fuse to generate an incomplete unit (Fig. 3). As *st530*, which has a small rearrangement in a conserved domain of just one unit, also results in a barely detectable mutant phenotype, it is possible that an aberrant unit in the deletion revertants is responsible for the residual mutant phenotype. Perhaps a precise fusion of units in a deletion event would result in a complete revertant (Figs 1 and 3).

Our results show that in large proteins containing repeated domains, large deletions can be consistent with normal, or nearly normal, functioning of the protein. Redundancy within the genome in the form of multigene families is common, and there are now several examples in which the null state of a gene is wild type^{13,14}. We suggest that *unc-22* illustrates the intragenic parallel to the intergenic phenomenon. Redundancy of functional units within a protein means that the absence of several functional units could be masked by the remaining units. Recently, weak alleles of some human disease loci have been identified that are associated with substantial genomic deficiencies⁴. Our results with *unc-22* provide a possible explanation for these observations. It will be interesting to see if the number of repeat units will vary across species in genes like *unc-22*.

In *unc-22*, perhaps of the repeated domains within the protein, we have a dramatic example of variability induced by transposon excision. Ds elements in maize also primarily excise imprecisely and also yield structurally and functionally altered proteins.

Transposon excision appears to be a potentially rich and evolutionarily significant source of genetic diversity for protein structure, and perhaps for allele heterogeneity in genetic disease.

We thank Mark Johnston and Daniela Gerhard for helpful discussions and for comments on the manuscript. This work was supported by the National Institutes of Health (R.H.W.).

Received 30 November 1987; accepted 11 January 1988.

1. Waterston, R. H., Thompson, J. N. & Brenner, S. *Dev. Biol.* **77**, 271–302 (1980).
2. Moerman, D. G., Benian, G. M. & Waterston, R. H. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2579–2583 (1986).
3. Moerman, D. G., Benian, G. M., Barstead, R. J., Schrieffer, L. & Waterston, R. H. *Genes Dev.* (in the press).
4. Coss, G. S. *et al.* *EMBO J.* **6**, 3277–3283 (1987).
5. Zengel, J. M. & Epstein, H. F. *Cell Motility* **1**, 73–97 (1980).
6. Greenwald, I. & Horvitz, H. R. *Genetics* **96**, 147–164 (1980).
7. Moerman, D. G. & Baillie, D. L. *Genetics* **91**, 95–104 (1979).
8. Moerman, D. G., Plurad, S. & Waterston, R. H. *Cell* **29**, 773–781 (1982).
9. Wessler, S. R., Baran, G., Varagona, M. & Dellaporta, S. L. *EMBO J.* **5**, 2427–2432 (1986).
10. Dooner, H. K. & Nelson, O. E. Jr *Proc. natn. Acad. Sci. U.S.A.* **76**, 2369–2371 (1979).
11. Schwartz-Sommer, Z., Gierl, A., Cuypers, H., Peterson, P. A. & Saedler, H. *EMBO J.* **5**, 591–597 (1985).
12. Starzyk, R. M., Webster, T. A. & Schimmel, P. *Science* **237**, 1614–1618 (1987).
13. Tatchell, K. D., Chaleff, D. T., Defeo-Jones, D. & Scolnick, E. M. *Nature* **309**, 523–527 (1984).
14. Landel, C. P., Krause, M., Waterston, R. H. & Hirsh, D. *J. molec. Biol.* **180**, 497–513 (1984).
15. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
16. Messing, J. *Meth. Enzym.* **101**, 20–78 (1983).
17. Staden, R. *Nucleic Acids Res.* **14**, 217–231 (1986).
18. Laemmli, U. K. *Nature* **227**, 680–685 (1976).
19. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4354 (1979).

Heterologous expression of a bacterial haemoglobin improves the growth properties of recombinant *Escherichia coli*

Chaitan Khosla & James E. Bailey*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA

Rational design of novel as well as improved cellular biocatalysts by genetic manipulation of cellular metabolism has recently attracted considerable interest. A wide range of bacteria have been genetically modified by integrating new enzymatic functions into their metabolic network^{1–7}. A central problem in the aerobic growth of any cell culture is the maintenance of dissolved oxygen (DO) concentrations above growth-limiting levels especially in high cell-density fermentations which are usually of a fed-batch type. The optimal rate of nutrient addition (and consequently the productivity) is ultimately limited by the rate at which cells can aerobically catabolize the carbon source without generating growth-inhibitory metabolites such as lactate and acetate^{8,9}. All approaches thus far have concentrated on improving the oxygen mass transfer rates by manipulating various environmental parameters. We¹⁰ have isolated the gene for a haemoglobin-like molecule, expressed by the aerobic bacterium *Vitreoscilla* in poorly-oxygenated environments^{11,12}, and expressed it in *Escherichia coli*. The recombinant cells contain enhanced haem as well as active haemoglobin, and they grow faster and to considerably greater cell densities than comparable plasmid-containing cells which do not express haemoglobin. This haemoglobin increases the rate of oxygen use, especially when dissolved oxygen is less than 5% of air saturation.

Recently, Wakabayashi *et al.*¹¹ reported the primary amino-acid sequence of a soluble dimeric haemoglobin found in the obligately aerobic bacterium, *Vitreoscilla*, and demonstrated significant similarity of this protein with known globin sequences. The synthesis of this protein increases severalfold in response

* To whom correspondence should be addressed.

Fig. 1 Comparison of growth dynamics of JM101 (▲—▲), JM101:pUC9 (■—■) and JM101:pRED2 (●—●) in fed-batch fermentations. Time courses of cell densities in batch (Fig. 1a, left) and fed-batch (Fig. 1b, right) modes of growth are shown.

Method. Cells were grown in a New Brunswick Microferm fermentor at $37 \pm 0.5^\circ\text{C}$ and a pH of 7 ± 0.05 with an initial working volume of 2.5 L. A constant air-flow rate of 4.5 l min^{-1} and agitator speed of 300 r.p.m. were maintained throughout each run. Foam control was achieved manually by intermittent addition of silicone AF60 antifoam emulsion as required. The batch medium and feed medium 1 of Table 2 in ref. 12 were used as batch and feed media compositions, respectively. For plasmid-containing cells, 100 mg l^{-1} ampicillin was added to the batch medium. At $t=0$, the fermentor was inoculated with a 100 ml overnight LB (1% Bactotryptone, 0.5% yeast extract, 1% NaCl) culture grown in a shake flask in each case. Cells were initially grown in batch mode until nutrient was exhausted, as indicated by a sudden increase in DO (Fig. 2). Thereafter, feeding was commenced (as indicated by vertical arrows, Fig. 1a; this is the initial point of each curve in Fig. 1b) at a constant rate of 10 ml h^{-1} and continued till the OD_{600} (optical density at 600 nm) reached a constant value.

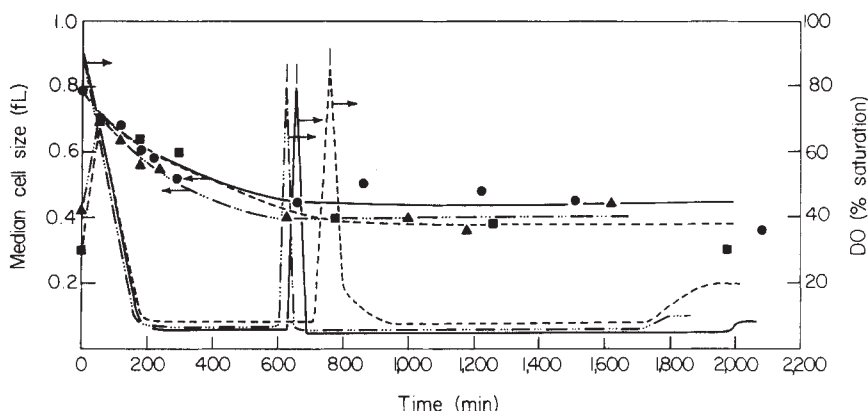
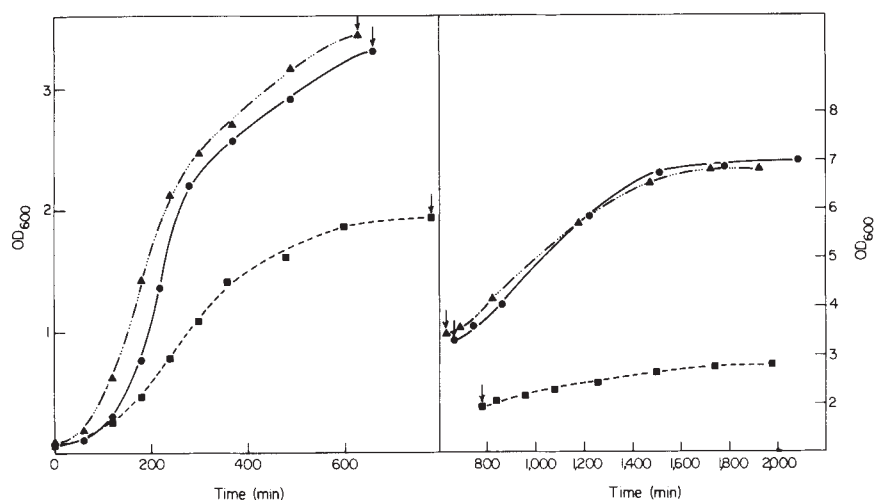


Fig. 2 Dissolved oxygen and median cell size profiles for JM101 (▲—▲), JM101:pUC9 (■—■) and JM101:pRED2 (●—●) in fed-batch fermentations.

Method. Details of the fermentation experiments are described in the caption to Fig. 1. DO was calibrated and recorded with reference to air saturation. The median cell size was determined for various samples (as indicated by the discrete points) on a Coulter counter, as described in ref. 18.

to a drop in DO levels, as evidenced by a rise in intracellular haem content¹². Although the exact molecular mechanism is unclear, several *in vitro* and *in vivo* biochemical studies on this protein suggest that it plays a haemoglobin-like role in the cell and is not a cytochrome as was initially suggested (for review see ref. 11).

Protohaem IX, the prosthetic group of all haemoglobins, is synthesized by wild-type *E. coli* as a component of its membrane-bound aerobic respiratory-chain proteins¹³. Mutants which overproduce this pigment without any apparent haem or cytochrome deficiency have been isolated¹⁴, suggesting that protohaem IX biosynthesis may be under repressive control, and hence amenable to amplification. Thus, bearing in mind the apparent *in vivo* role of the *Vitreoscilla* haemoglobin as well as the importance of *E. coli* hosts for high cell-density recombinant fermentations, we attempted functional expression of this gene in *E. coli*. Using mixed oligonucleotide probes, we have isolated and cloned into pUC19 a *Hind*III fragment from a *Vitreoscilla* genomic library that codes for the haemoglobin gene¹⁰. The gene expressed a polypeptide of the correct relative molecular mass as a major cellular protein in *E. coli* when grown under culture-tube or shake-flask conditions. This expression is apparently regulated by endogenous *Vitreoscilla* upstream sequences. More significantly, greater than fivefold overproduction of intracellular haem was spectrophotometrically recorded. Such elevated haem levels impart a distinct reddish tint to these recombinant cell pellets compared to controls. The haemoglobin activity of these cells was correspondingly fivefold higher than controls as measured by the carbon monoxide difference spectrum method of Webster and Liu¹⁵.

To determine whether the functional expression of this protein had any metabolic effect(s) on the host, we compared the growth properties of this recombinant strain (JM101:pRED2, ref. 10) with similar plasmid-containing (JM101:pUC9) and plasmid-free (JM101) strains under typical fed-batch fermentation conditions (Fig. 1). In each case the fed-batch phase of the experiment was conducted under oxygen-limited conditions⁹. The growth parameters corresponding to the data shown in Fig. 1 are summarized in Table 1. In all cases, the correlation between the optical density (OD) and the dry-cell weight was linear, as determined by intermittent dry-cell weight measurements, thus

Table 1 The growth parameters of JM101, JM101pUC9 and JM101:pRED2

	JM101	JM101:pUC9	JM101:pRED2
Batch log-phase growth rate* (h^{-1})	0.95	0.73	0.95
Batch stat-phase dry cell mass† (g l^{-1})	2.6	1.6	2.6
Fed-batch log-phase growth rate* (h^{-1})	0.056	0.033	0.064
Final dry cell mass‡ (g l^{-1})	5.8	2.8	5.9

* As calculated from the data presented in Fig. 1.

† Corresponding to vertical arrow positions in Fig. 1a.

‡ Dry-cell mass was determined for 10 ml aliquots taken directly from the fermentor. Cells were centrifuged, washed once and resuspended in deionized water, all below 4°C . Samples were dried overnight in a 100°C oven.

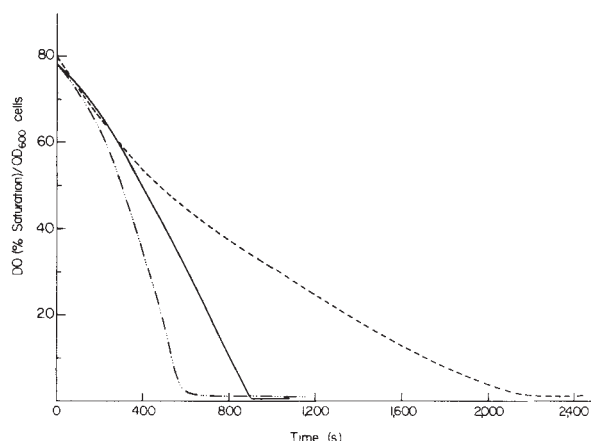


Fig. 3 Respiratory behaviour of JM101 (▲—····—▲), JM101:pUC9 (■---■) and JM101:pRED2 (●—●).

Method. 1 ml stationary-phase cells (from samples indicated by vertical arrows in Fig. 1a) were centrifuged and resuspended in 25 μ l 100 mM phosphate buffer (pH 7). This was then injected into fresh medium prewarmed to 37°C in a Gilson respirometer (~2 mL). The above curves represent the time course of DO recordings normalized to a unit OD₆₀₀ basis.

justifying the use of OD as a measure of growth. The JM101:pRED2 strain significantly outgrew the control recombinant strain JM101:pUC9 throughout the experiment and also exhibited a higher specific growth rate than the plasmid-free strain JM101 during the fed-batch phase. Figure 3 illustrates important differences between haemoglobin-containing and haemoglobin-free cells with respect to their oxygen use kinetics, particularly at low DO concentrations. It should be pointed out that though the measured DO levels in the fed-batch phases are around 5% (Fig. 2), spatial microheterogeneity in bioreactors can result in pockets of considerably lower DO levels and poor mixing^{16,17}; these effects are of greater significance at higher cell densities and in larger reactors.

Haemoglobin activity of JM101:pRED2 samples withdrawn from the fermentor could not be spectrophotometrically detected by the above-mentioned method¹⁵ using whole-cell suspensions; these were considerably more dilute than recommended due to sample-size limitations. However small quantities of the polypeptide were observed relative to controls on a silver-stained sodium dodecyl sulphate (SDS) polyacrylamide gel. (This analysis required considerable darkening of the gel such that it is unrepresentable. We are now working on the development of a more sensitive assay for the active protein.) This led us to perform a series of shake-flask experiments to study the oxygen-dependent regulation of this gene (C. Khosla and J.E.B., manuscript in preparation). Two observations from those experiments are relevant in this context. First, expression of the haemoglobin gene in *E. coli* is under oxygen-dependent regulation in a manner similar to that in *Vitreoscilla*. The major difference seems to be that at very low DO concentrations, as are often found in culture tubes and flasks, the gene in recombinant *E. coli* is overexpressed to form inclusion bodies. This is consistent with a larger median cell size observed for the JM101:pRED2 inoculum (which came from the stationary phase of a shake-flask preculture) compared to those of growing cells in the fermentor (Fig. 2). Second, presence of the cloned haemoglobin gene has provided growth enhancement in terms of dense-culture growth rates as well as final cell densities under all conditions studied so far. These experiments suggest a possible utility for this haemoglobin or similar molecules in any aerobically grown cell lines in which the haemoglobin can be functionally expressed and appropriately localized.

The exact mechanism underlying these phenomena is unknown. The haemoglobin protein probably does not function only as an oxygen storage trap; it is probably involved in enhancing the oxygen uptake rate of the membrane-bound aerobic respiratory apparatus, a role which the haemoglobin might conceivably play even in low concentrations. As the regulatory mechanism of this gene is functional in *E. coli*, its genetic dissection should be relatively straightforward.

This work has been supported by the Energy Conversion and Utilization Technology (ECUT) program of the United States Department of Energy. We thank Dr Webster for providing the *Vitreoscilla* strain.

Received 7 December 1987; accepted 4 January 1988.

1. Reineke, W. & Knackmuss, H. J. *Nature* **277**, 385–386 (1979).
2. Ensley, B. D. *et al. Science* **222**, 167–169 (1983).
3. Anderson, S. *et al. Science* **230**, 144–149 (1985).
4. Hopwood, D. A. *et al. Nature* **314**, 642–644 (1985).
5. Mermod, N., Harayama, S. & Timmis, K. N. *Biotechnology* **4**, 321–324 (1986).
6. Stanzak, R., Matsushima, P., Baltz, R. H. & Rao, R. N. *Biotechnology* **4**, 229–232 (1986).
7. Windass, J. D. *et al. Nature* **287**, 396–401 (1980).
8. Zabriskie, D. W. & Arcuri, E. J. *Enzym. microb. Technol.* **8**, 706–717 (1986).
9. Tsai, L. B. *et al. J. ind. Microbiol.* **2**, 181–187 (1987).
10. Khosla, C. S. & Bailey, J. E. *Molec. gen. Genet.* (submitted).
11. Wakabayashi, S., Matsubara, H. & Webster, D. A. *Nature* **322**, 481–483 (1986).
12. Boerman, S. J. & Webster, D. A. *J. gen. appl. Microbiol.* **28**, 35–43 (1982).
13. Anraku, Y. & Gennis, R. B. *Trends biochem. Sci.* **12**, 262–266 (1987).
14. Cox, R. & Charles, H. P. *J. Bact.* **113**, 122–132 (1973).
15. Webster, D. A. & Liu, C. Y. *J. biol. Chem.* **249**, 4257–4260 (1974).
16. Oosterhuis, N. M. G. & Kosen, N. W. F. *Biotechnol. Bioengng.* **26**, 546–550 (1984).
17. Vardar, F. *Process Biochem.* **18**, 21–23 (1983).
18. Wittrup, K. D., Mann, M., Fenton, D., Tsai, L. & Bailey, J. E. *Biotechnology* (submitted).

Expression of the putative Duchenne muscular dystrophy gene in differentiated myogenic cell cultures and in the brain

Uri Nudel, Kenneth Robzyk & David Yaffe

Department of Cell Biology, The Weizmann Institute of Science, Rehovot 76100, Israel

Duchenne muscular dystrophy (DMD), a sex-linked degenerative disorder of the muscle, is one of the most common lethal genetic diseases in man. It affects about one male in 3,500, with an estimated one-third of cases being caused by new mutations. A less severe disease, Becker's muscular dystrophy (BMD), maps to the same chromosomal locus and is most probably an allelic form of DMD. Both diseases are sometimes associated with various degrees of mental retardation; the molecular basis of these phenotypes is unknown (for review, see ref. 1). The giant DMD gene spans approximately 2,000 kilobases (kb) (0.05% of the human genome) and encodes a 14-kb mRNA^{2,3}. The tissue-specificity of its expression has not been precisely determined. Monaco *et al.*², using Northern blots, reported expression of the gene in human fetal skeletal muscle and small intestine but not in human fetal brain, or in human cultured myoblasts and transformed B and T cells. More recently, expression was detected in mouse skeletal and cardiac muscle, but not in mouse brain⁴. Here we show, using a ribonuclease protection assay, that the DMD gene is developmentally regulated in rat and mouse myogenic cell cultures, and that it is expressed in rat and mouse striated muscle, in mouse smooth muscle and in rat, mouse and rabbit brain. We could not detect transcripts in other non-muscle tissues.

Using two synthetic oligonucleotides complementary to the potential coding region of human clone 87-25 (ref. 2) as hybridization probes, we isolated a homologous rat genomic DNA clone. A 500-base pair (bp) *EcoRI*–*PvuII* DNA fragment was sub-

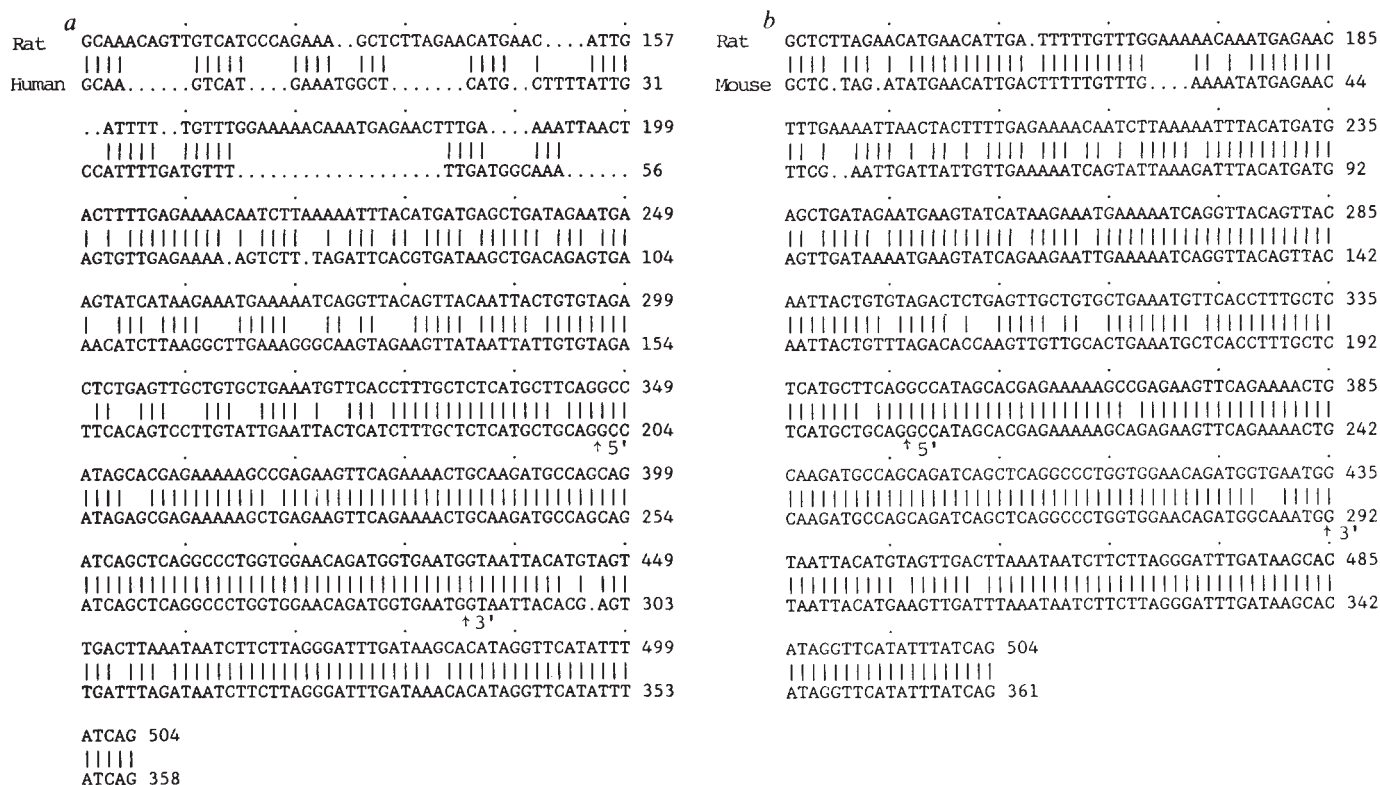


Fig. 1 Sequence homology between rat, mouse and human DMD gene segments. *a*, Comparison of rat (upper) and human² (lower) sequences; *b*, comparison of rat (upper) and mouse² (lower) sequences. (|) Indicates sequence homology. Arrows indicate the potential 5' and 3' splice sites as suggested by Monaco *et al.*².

cloned into the vector Gemini3 (Promega) and was sequenced (pMD1). As shown in Fig. 1, the regions which are highly conserved in the human and mouse genes are also conserved in rat. The homology between the rat and mouse sequences upstream from the potential coding region is greater than the homology between human and either rodent. Interestingly, the sequences downstream from the potential exon are highly conserved in all three species, suggesting that this region, which probably does not encode amino acids, has some other functional role.

The 500-bp fragment containing the highly conserved sequences was used to probe Southern blots of *Hind*III- or *Bam*HI-digested DNA isolated from hybrid cell lines containing overlapping subsets of mouse chromosomes in a Chinese hamster genome background. The probe hybridized strongly to a single mouse and Chinese hamster DNA fragment, indicating that this fragment is present in a single copy in these species. According to the pattern of hybridization, the mouse fragment was assigned to the X chromosome (data not shown).

The tissue specificity of expression of the DMD gene in mouse was studied using an RNA probe transcribed from pMD1, and the RNase protection assay⁵. An RNA fragment of approximately 85 bp was protected by total RNA prepared from mouse striated muscle (leg skeletal muscle and cardiac muscle), smooth muscle (stomach; not shown) and, very interestingly, from mouse brain. Protection was not detected with RNA from spleen (not shown), liver or kidney (Fig. 2). An RNA fragment, approximately 90 nucleotides long, was protected by RNA from rat skeletal muscle, cardiac muscle and brain. The same RNA fragment was protected by skeletal muscle (Fig. 2) and cardiac RNA preparations from rabbits (not shown). The size of the labelled RNA fragment protected by the rat derived RNA is consistent with the size of an exon predicted by Monaco *et al.*² and later verified by Hoffman *et al.*⁴. Protection was not observed

when the opposite strand was used as a labelled RNA probe (data not shown).

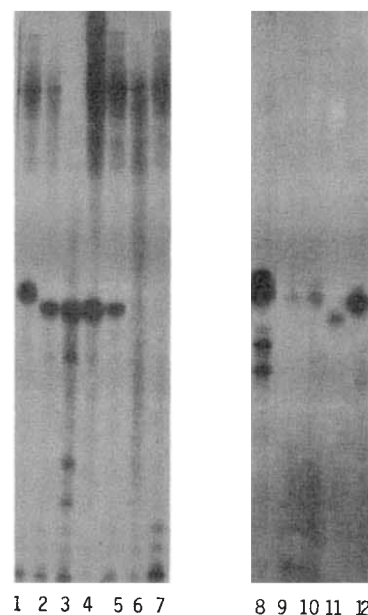
To determine the stage of cellular differentiation at which the DMD gene is expressed, we analysed the expression of the DMD gene in rat primary skeletal muscle cell cultures and in established myogenic cell lines. The RNA preparations from primary muscle cultures were also assayed for the expression of the fast muscle myosin light chain 2 (MLC2) gene. As shown in Fig. 3, DMD gene transcripts were not detected or barely detected in undifferentiated myogenic cultures consisting of mononucleated cells prepared from newborn rat skeletal muscle. The DMD mRNA accumulates, as does the MLC2 mRNA, during the phase of rapid cell fusion and formation of multinucleated myotubes^{6,7}.

A relatively large quantity of DMD transcripts was detected in RNA preparations from differentiated myogenic MAG6 cells (a myogenic cell line derived from mouse embryonal carcinoma cells; D. Shinar *et al.*, in preparation). As in primary cultures, the DMD transcript was barely detectable in RNA preparations from undifferentiated mononucleated MAG6 cells. This minimal signal may be accounted for by the presence in such cultures of a very small percentage of multinucleated cells (Fig. 3). Examination of additional myogenic cell lines revealed a great variability in the level of expression of the gene (not shown). In all myogenic cell lines examined in which expression of the gene was detectable, appreciable amounts of DMD mRNA were found only in cultures which had been induced to differentiate.

The 500-bp probe was hybridized to Northern blots of total RNA isolated from rat skeletal muscle, rat brain, mouse liver, kidney and spleen, and differentiated L8 and MAG6 cells. The probe hybridized to a discrete band corresponding to an RNA species of approximately 14 kb (and a smear below) in muscle and MAG6 RNA. A faint band was also observed with brain

Fig. 2 DMD gene expression in various tissues. A labelled RNA probe transcribed from pMD1 was protected from RNase digestion by RNA samples from the following tissues: 1, Rat skeletal muscle; 2, newborn mouse skeletal muscle; 3, mouse cardiac muscle; 4, mouse skeletal muscle; 5, mouse brain; 6, mouse liver; 7, mouse kidney; 8, rabbit skeletal muscle; 9, newborn rat brain; 10, adult rat brain; 11, LC4 differentiated cultures (a mouse myogenic cell line, unpublished); 12, rat skeletal muscle. Mouse RNA protected a shorter stretch of the rat-derived probe, most probably because of the divergence between rat and mouse sequences at positions 429 and 430 of the rat sequence, five and six nucleotides upstream from the potential 3' splice site (Fig. 1).

Methods. RNA was prepared from various tissues using the lithium chloride/urea extraction method¹⁵. Total RNA (5 µg) from various tissues were hybridized overnight with a uridine P³²-labelled RNA probe (300,000 c.p.m., specific activity 2×10^8 c.p.m. µg⁻¹) at 50 °C in 20 µl of buffer containing 80% formamide, 40 mM PIPES, 0.4 M NaCl and 1 mM EDTA. The samples were then incubated for 60 min at 37 °C with 8 µg RNase A and 0.4 µg RNase T1 in 200 µl buffer containing 10 mM Tris pH 7.5, 0.3 M NaCl and 5 mM EDTA. The RNA was precipitated, dissolved in a sample buffer containing 80% formamide in TBE buffer (89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA) and then electrophoresed on a 6% polyacrylamide-urea sequencing gel and fluorographed.



Probe: DMD MLC2 DMD

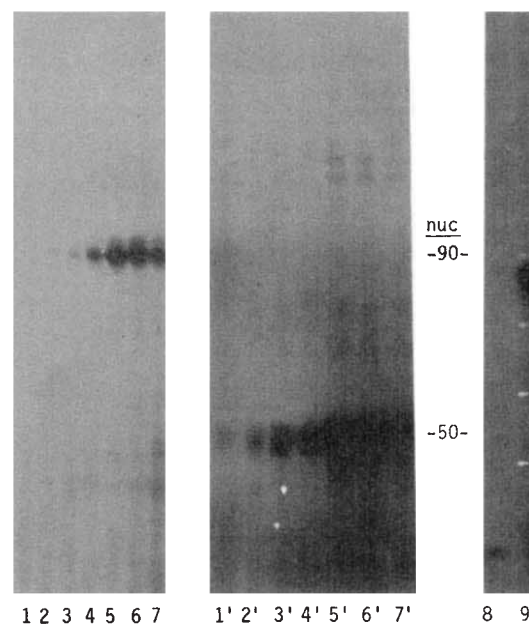


Fig. 3 Accumulation of DMD RNA during differentiation of myogenic cells. Mononuclear cells from the leg muscle of newborn rats were cultured as described previously¹⁶. Differentiation promoting medium (S-medium) was added 48 h after plating (time zero). RNA was then extracted at the following timepoints: zero (lanes 1 and 1'); 14 h (2, 2'); 17.5 h (3, 3'); 23 h (4, 4'); 38 h (5, 5'); 43 h (6, 6') and 64 h (7, 7'). Rapid cell fusion began at approximately 18 h after the change to S-medium. Total RNA (5 µg) was assayed for the presence of DMD RNA sequences as described in the legend to Fig. 2 (lanes 1-7). These preparations were also assayed for MLC2 mRNA, using the RNase protection method and a probe complementary to the MLC2 mRNA 5' region (lanes 1'-7'). Additional RNA preparations assayed with the DMD probe were from undifferentiated (lane 8) and differentiated (g) MAG6 cells (a myogenic cell line derived from mouse embryonal carcinoma; Shinar *et al.*, in preparation).

RNA. No hybridization was detected with RNA preparations from the other non-muscle tissues and myogenic L8 cells⁸ (data not shown).

Our results demonstrate that the DMD gene is expressed in a tissue-specific manner, and that it is developmentally regulated in myogenic cells. The detection of DMD transcripts in brain is of special significance as it may explain the association of DMD and BMD with various degrees of mental retardation. Moreover, since the gene is expressed in both muscle and brain, the data do not exclude the suggestion made by several investigators that the disease might be of neurological or neuromuscular aetiology^{1,9,10}.

The high degree of homology between the sequenced fragment of the rat gene and the published sequence from the human DMD gene, the Southern blot hybridizations in which a single band assigned to the mouse X chromosome is detected, the

Northern blot analysis in which a 14-kb band is observed, and the sizes of the riboprobe fragments protected by rat and mouse RNA all indicate that pMD1 is indeed a specific probe for the rat gene homologous to the putative human DMD gene and for its transcript.

The failure of other investigators^{2,4} to detect the DMD transcript in the brain, using a probe for the homologous sequence in the human gene, may be due to the use of different and/or less sensitive probes and methods. We also found that the RNA blotting and hybridization assay (Northern), used by these authors, is much more susceptible to false negative results due to partial degradation of the high-molecular-weight RNA than the RNAase protection assay method used here which is based on the protection of a relatively short RNA probe. Alternatively, the discrepancy between our results and those described in refs 2 and 4 could be due to alternative splicing of the DMD gene

transcript in the muscle and in the brain (and the use of different probes), or perhaps to the lack of poly(A) tail in the brain DMD mRNA (and the use of total versus polyadenylated RNA). It should also be noted that the level of expression of the gene in the brain is lower than its expression in skeletal muscle (about $\frac{1}{3}$ – $\frac{1}{10}$).

The mode of expression of the gene in myogenic cells strongly suggests that the gene encodes a protein which is involved in terminal differentiation of muscle fibers. The time course of DMD mRNA accumulation closely correlates with that of mRNAs coding for many muscle-specific contractile proteins and enzymes involved in muscle contraction. These data, together with the reported RNA size, are consistent with the suggestion that the gene encodes a very large protein associated with the contractile apparatus. However, other differentiation-associated proteins (such as cell membrane proteins) cannot be excluded.

In a very recent publication, Hoffman *et al.*¹¹ reported the identification of a protein of relative molecular mass 400,000 which reacts with antibodies raised against fusion proteins containing a polypeptide encoded by the DMD gene. The data suggest that this protein is associated with the triadic structures which are located in the A-I junctions of muscle sarcomeres and which are involved in the coupling between membrane excitation and contraction of the myofibrils^{12,13}. Hammonds¹⁴ showed that the amino-terminal sequence of the predicted DMD protein has significant homology with a sequence of the actin-binding domains of α -actinin. These data are consistent with our results showing that the DMD mRNA accumulates during terminal differentiation of myogenic cells.

The apparent variability in the level of expression of this gene in various myogenic cell lines provides a convenient cell culture system for studying the gene product(s) and its biological function.

We thank Ms Z. Levy and S. Neuman for technical assistance; Dr I. Schechter for the rat gene library; P. Einat, Y. Mayer and D. Shinar for RNA samples; Doron Shinar for the MAG6 cells, and Dr H. Czosnek for the Southern blots of the *EcoRI* digested DNA isolated from mouse Chinese hamster hybrid cell lines.

This work was supported by the Muscular Dystrophy Association, New York, and the Leo and Julia Forchheimer Center for Molecular Genetics at the Weizmann Institute of Science.

Received 30 November 1987; accepted 7 January 1988.

1. Moser, H. *Hum. Genet.* **66**, 17–40 (1984).
2. Monaco, A. P. *et al. Nature* **323**, 646–650 (1986).
3. Monaco, A. P. & Kunkel, L. M. *Trends Genet.* **3**, 33–37 (1987).
4. Hoffman, E. P., Monaco, A. P., Feener, C. C. & Kunkel, L. M. *Science* **238**, 347–350 (1987).
5. Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035–7056 (1984).
6. Yabionka, Z. & Yaffe, D. *Differentiation* **8**, 133–143 (1977).
7. Shani, M. *et al. Dev Biol.* **86**, 483–492 (1981).
8. Yaffe, D. & Saxel, O. *Differentiation* **7**, 159–166 (1977).
9. Sica, R. E. P. & McComas, A. J. *Can. J. Neurol. Sci.* **5**, 189–197 (1978).
10. Rowland, L. P. *Archs Neurol.* **33**, 315–321 (1976).
11. Hoffman, E. P., Brown, R. H. & Kunkel, L. M. *Cell* (in the press).
12. Hoffman, E. P., Knudson, C. M., Campbell, K. P. & Kunkel, L. M. *Nature* **330**, 754–758 (1987).
13. Slater, C. R. *Nature* **330**, 693–694 (1987).
14. Hammonds, R. G. *Cell* **51**, 1 (1987).
15. Auffray, C., Nageotta, R., Chambraud, B. & Rougeon, F. *Nucleic Acids Res.* **8**, 1231–1241 (1980).
16. Yaffe, D. in: *Tissue Culture, Methods and Applications* (eds Kruse, P. F. & Patterson, M. K.) 106–114 (Academic, New York, 1973).

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI
Article
Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem
☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____
 Title _____
 Institution/Company _____
 Department _____
 Address _____
 City _____ State _____ Zip _____
 Phone (____) _____

Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

Capillary zone electrophoresis-MS

R.D. Smith and H.R. Udseth

Coupling the separation capabilities of capillary zone electrophoresis with the analytical specificity of mass spectrometry yields a tool for detecting levels of substances down to 10 attomoles.

CAPILLARY zone electrophoresis (CZE) is a form of free-zone electrophoresis capable of fast high-resolution separations. As initially described by Jorgenson and coworkers^{1,2}, CZE typically is conducted in fused silica tubing capillaries of between 25 and 150 μm i.d., using voltages of between 10 and 40 kV. Separations result from differences in the electrophoretic mobilities of analytes, although the inclusion of selected components in the CZE buffer medium can also be used to modify the migration process. CZE separations generally occur superimposed upon a bulk electroosmotic flow which originates in the electrical double layer at the capillary-buffer interface.

CZE is therefore an elution technique which resembles chromatography and is readily amenable to automation. The flat flow profiles from electroosmosis and the suppression of convection in the small capillaries allow extremely large numbers of theoretical plates (10^5 to 10^7) to be obtained readily^{1,3}. CZE is best viewed as a highly efficient analytical micro-separation method, with sample volumes on the order of 1 to 100 nl. These characteristics make CZE well suited for the analysis of biological samples, but ideally require the use of detection methods that are sensitive, selective and universally applicable.

Conceptually, the combination of the separation capability of CZE and the analytical specificity of mass spectrometry (MS) is a nearly ideal marriage of instrumental techniques. But the technical requirement that both ends of the capillary be immersed in buffer reservoirs of greatly differing voltages — a situation that wedded CZE to such “on-line” methods as UV and fluorescence detection¹ — made the straightforward union of CZE and MS difficult. To overcome this, at the CZE-MS interface an electrical contact was made with the eluting liquid at the capillary terminus, so that the liquid formed an electrospray at atmospheric pressure in the presence of a high electrical field⁴. The CZE-MS instrumentation developed at our laboratory and the details of one version of the interface are schematically illustrated in Fig. 1.

The CZE-MS interface is based upon an electrospray ionization (ESI) process that allows ion production from the end of the CZE capillary at atmospheric pressure by an electrically-induced nebulization process⁴. The highly charged droplets are exposed to a countercurrent flow of dry

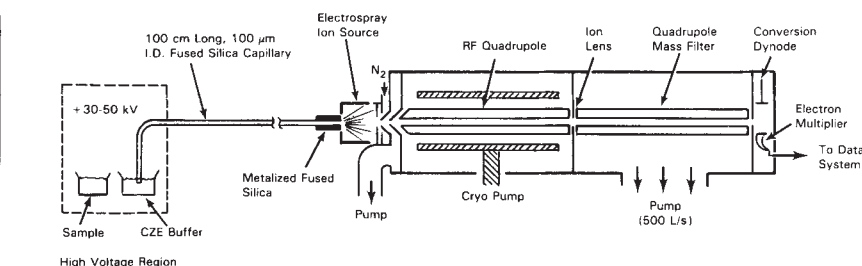


Fig. 1 Schematic of CZE-MS instrumentation. The CZE separation and ESI occur at atmospheric pressure. The countercurrent flow of warm nitrogen allows only nitrogen and ions produced by ESI to enter the sampling orifice of the mass spectrometer.

nitrogen which facilitates evaporation and causes the droplets to reach their charge limit. Subsequent evaporation results in droplet breakup, producing smaller droplets which proceed along a similar course until the droplets are sufficiently small that direct field-assisted ion desorption (or ion evaporation) can occur from the droplet surface⁵. The precise sequence of events, the competition between the above processes, and the dependence of the process upon droplet size are not fully understood. The role of liquid composition upon ion desorption is also largely unexplored. However, the net result of the ESI process is the desorption of analyte ions by a gentle process which does not require heating. The 3 to 6 kV potential gradient used for ESI can produce positive or negative ions, depending upon the field polarity. In our current instrument configuration, we have utilized a three-stage vacuum system incorporating a mechanically pumped first stage and a cryogenically pumped second stage. The second stage contains a radio frequency quadrupole lens which transmits ions efficiently to the quadrupole mass spectrometer in the third stage.

The CZE-MS electrospray interface is effective because the CZE buffer ionic strength and flow rate overlap with those required for electrospray ionization⁴. Electrospray ionization was first explored for mass spectrometry by Dole⁶, and more recently Fenn and coworkers reported on the potential for interfacing liquid chromatography to mass spectrometry⁷. The ESI process is amenable to a wide range of polar or ionic hydrophilic materials, and experience to date suggests it is most effective for substances which may exist as charged species in solution, where conventional mass spectrometric ionization techniques are generally inadequate. ESI also can provide exceptional detection limits^{4,7}. A current limitation of

the CZE-MS interface is the limited understanding of ESI and the lack of systematic methods for obtaining mass spectra through the manipulation of buffer composition. However, the study of “desorption” ionization methods is currently a major research direction in mass spectrometry and rapid advances in both the understanding and manipulation of the ESI process are anticipated⁸.

The use of CZE-MS has been limited because it is not readily adaptable to “conventional” mass spectrometers, and it requires an instrumental configuration similar to that required for atmospheric pressure ionization (API). Methods that have been developed to make CZE more suitable for use with conventional mass spectrometers appear to have severe limitations. Having eluting liquid droplets enter through an orifice or capillary for introduction into a conventional ion source, for example, is undesirable because the eluent would be transported inefficiently and only compounds having sufficient vapour pressure at ion source conditions could be analysed. To avoid these problems, the CZE separation capillary could be incorporated into a “flowing” fast atom bombardment (FAB) interface, or ions eluting from the capillary could be desorbed directly by electrohydrodynamic ionization⁹. Both approaches present sensitivity limitations and are adversely affected by the pressure drop across the capillary caused by the vacuum requirements of the mass spectrometer, and so might be restricted to use with highly viscous CZE media. ESI is effective at atmospheric pressure, avoiding a pressure drop across the capillary, and eliminating a contribution of laminar flow. (The parabolic profile of laminar flow results in a substantial degradation of CZE separation efficiency.) While an ideal interface to “conventional” mass spectrometers appears elusive, the ESI

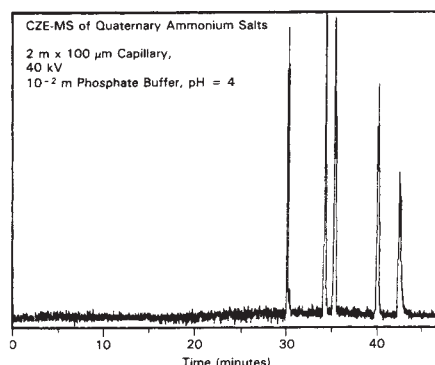


Fig. 2 Total ion electropherogram obtained for a mixture of quaternary ammonium salts by CZE-MS. The ESI positive ion spectra consist of only the corresponding quaternary ammonium cations.

approach offers numerous advantages, and recent developments in our laboratory continue to enhance its ease of use and performance. In addition, there is a growing recognition of the advantages of API sources in mass spectrometry. Several API instruments are commercially available, and interfaces with both supercritical fluid chromatography and liquid chromatography appear feasible.

Figure 2 shows a CZE-MS total ion electropherogram obtained for a mixture of quaternary ammonium salts. The number of theoretical plates in the separation ranges from 160,000 for the trimethylphenylammonium ions to 330,000 for tetrabutylammonium ions (the last peak), an order of magnitude higher than that obtainable by liquid chromatography in a similar time. In this case sample sizes were 300 fmol, but detection limits of 10 amol have been obtained by single ion detection.

The CZE-MS combination is complementary as CZE separations are based on the charge state of the analyte in solution, and ESI seems to function most effectively for ionic species. ESI spectra are often dominated by molecular ions and multiply-charged species can be significant contributions to mass spectra. Results for polypeptides suggest that such "multiply-charged spectra" obtained under some conditions may be useful for compound identification¹⁰. □

Richard D. Smith and Harold R. Udseth are at the Chemical Sciences Department, Battelle Pacific Northwest Laboratory, Richland, Washington 99352, USA. For more information, fill in reader service number 100.

1. Jorgenson, J.W. & Lukacs, K.D. *Science* **222**, 266-272 (1984).
2. Jorgenson, J.W. & Lukacs, K.D. *Analyt. Chem.* **53**, 1298-1302 (1981).
3. Lauer, H.H. & McManigill, D. *Analyt. Chem.* **58**, 166-170 (1986).
4. Olivares, J.A., Nguyen, N.T., Yonker, C.R. & Smith, R.D. *Analyt. Chem.* **59**, 1230-1232 (1987).
5. Thomson, B.A. & Iribarne, J.V. *J. chem. Phys.* **71**, 4451-4463 (1979).
6. Dole, M. *et al. J. chem. Phys.* **49**, 2240-2249 (1968).
7. Whitehouse, C.M., Dreyer, R.N., Yamashita, M., & Fenn, J.B. *Analyt. Chem.* **57**, 675-679 (1985).
8. Vestal, M.L. *Mass Spectrom. Rev.* **2**, 447-480 (1986).
9. Cook, K.D. *Mass Spectrom. Rev.* **5**, 467-519 (1986).
10. Smith, R.D., Olivares, J.A., Nguyen, H.T. & Udseth, H.R. *Analyt. Chem.*, in the press.

Spectroscopy at Pittcon

Next week, thousands of analytical chemists will flock to New Orleans, Louisiana for the annual Pittsburgh Conference. Instrumentation for spectroscopy will dominate the show; a few selections are outlined below.

In booths 1532-1543, Philips will be exhibiting its model PV9550 energy-dispersive **X-ray spectrometer** system (*Reader Service No. 101*). Philips says the \$60,000 (US) instrument is designed for busy laboratories where single-element techniques such as atomic absorption spectroscopy are too slow and labour-intensive



X-ray spectrometry from Philips.

and where more sophisticated techniques such as inductively-coupled plasma mass spectroscopy are too expensive and unwarranted. The PV9550 permits the identification and measurement of all elements from sodium to uranium, down to a few parts per million. The spectrometer can perform non-destructive testing for the multi-element analysis of solids, thin films, powders or liquids, with minimal sample preparation. The PV9550's specimen chamber will accommodate samples ranging in size from 140 mm in diameter and 100 mm high, down to particles measuring a fraction of a millimeter across. Samples can be changed quickly: Philips says an air-path exchange can be performed in seconds, and a vacuum-path operation in less than a minute.

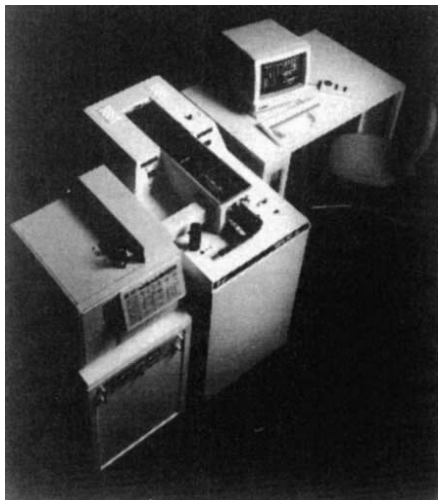
VG Elemental, a member of the VG Instruments Group exhibiting in booths 434-543, has a range of instruments for the **trace element analysis** of solids, liquids and gases (*Reader Service No. 102*). The £350,000 (UK) VG9000 glow discharge mass spectrometer is geared toward the quantitation of trace elements from the entire periodic table — including gallium, indium, silicon and gallium arsenide — down to several parts per billion. For more precise quantitation, the £160,000 (UK) VG PlasmaQuad inductively-coupled mass spectrometer performs a full multi-element analysis of liquids below the parts per thousand million level within minutes, says VG Elemental. A new option for use with the VG PlasmaQuad, the

VG Laserlab, provides rapid analyses by the direct laser ablation of solid samples.

In booth 2546, the French company G.I.P. Instrumentation et Spectrometrie will be introducing its advanced prototype multichannel **microspectrofluorometer** with laser excitation into the United States (*Reader Service No. 103*). G.I.P. says the microspectrofluorometer has potential applications in the study of intercellular fluorescence in single cells. The prototype consists of a UV He-Cd Omnichrome laser, a microscope with reflected light fluorescence attachment and UV optics, a triple spectrograph with fully reflective coupling optics between the microscope and the spectrograph, a multichannel intensified array detector, and a microcomputer for controlling the wavelength setting, positioning the laser on the sample and acquiring and analysing data.

Mass spectrometry

In booth 1234, Finnigan MAT will be unveiling its new high-performance, **single-stage quadrupole mass spectrometer**, the model SSQ 70 (*Reader Service No. 104*). The SSQ 70 is based on the design of Finnigan MAT's triple-stage quadrupole MS/MS system, the model TSQ 70, and can be upgraded to the TSQ 70 as required. Finnigan MAT says the hyperbolic quadrupole analyser gives the SSQ 70 a mass range of up to 4,000 AMU, and that a high voltage conversion dynode multiplier and an ion detection system with automatically adjusting bandwidth



Finnigan MAT's single-stage quadrupole MS can be upgraded to a triple-stage instrument as your needs grow.

guarantee maximum sensitivity at any scan rate and mass. Optional accessories for the SSQ 70 include a capillary GC interface, FAB probe, thermospray LC/MS interface, a direct exposure probe, a supercritical fluid chromatography interface and a solids probe.

Jeol will be exhibiting a range of products for spectroscopy in booths 141–155. Among the offerings from Jeol is its model GSX NMR spectrometer (*Reader Service No. 105*). The £130,000–350,000 (UK) GSX NMR spectrometer has an acquisition processor that can handle five million words of 32-bit memory. Matrix



Jeol's NMR spectrometer can be used with any one of 50 probes.

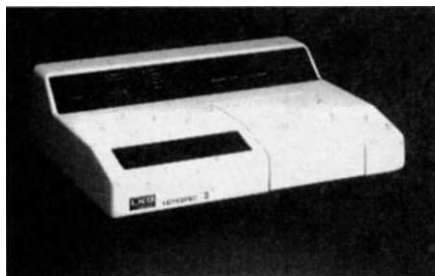
shims are fitted as standard, enabling the operator to compute and correct fifth- and sixth-order gradients on installation, using a shim robot system. The system's software simplifies shimming, phasing and integration to one-word commands. Nearly 50 probes are available for use with the GSX NMR spectrometer, including single nucleus probes, probes for dual C-H or dual F-H operation and probes tunable over a range of nuclei. A proton probe with a signal-to-noise ratio of 250:1 is also offered. The system accommodates 168 Mbyte Winchester disks upgradable to 825 Mbytes, and can be supported by an array processor.

Kratos Analytical, in booths 2617–2724, has developed a modular approach to big-ticket spectroscopy — the Concept system (*Reader Service No. 106*). With the Concept mass spectrometer system, researchers can purchase instrumentation

to suit their current needs, without worrying that the equipment will be insufficient to grow with them as their needs become more sophisticated. Concept is built around five fundamental units. Two operating consoles — one with fully automated controls and the other with additional manual peak matching — are available. Users can choose between a 2,000 AMU S magnetic sector mass analyser and a 10,000 AMU H mass analyser, and a quadrupole analyser can be added to either one. Sources available include EI, CI, EI-CI, FAB and LSIMS. Kratos Analytical says the high angular acceptance of the instrument, 0.012 rad, together with its long slits and short ion path give the Concept excellent transmission characteristics.

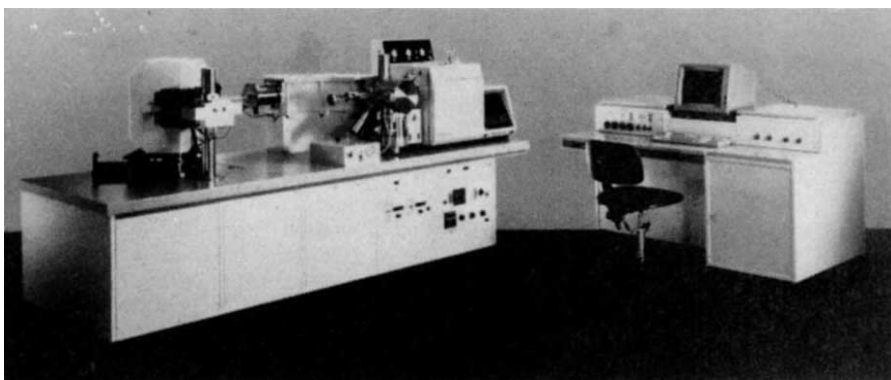
Spectrophotometry

Pharmacia LKB's Ultrospec II 4050 UV-visible spectrophotometer can be used to calculate the absorbance ratio between any two selected wavelengths — a plus in applications such as the determination of the purity and concentration of protein and DNA or RNA samples (*Reader Service No. 107*). The spectrophotometer's new microcuvette and holder



The Ultrospec II 4050 UV-visible spectrophotometer performs ratio calculations.

allows as little as 50 µl of sample to be analysed with virtually complete sample recovery, says Pharmacia LKB. The black walls of the microcuvette eliminate side wall transmission ensuring measurement accuracy, and the V-shaped sides facilitate dispensing. The cell holder is spring-loaded in two planes for better reproducibility of results. Pharmacia LKB will be in booths 223–235 and 322–334.



The Concept modular mass spectrometer from Kratos Analytical suits needs from the simple to the sophisticated.

The French company Secomam will be exhibiting an updated version of its popular A.B.C. programmable spectrophotometer in booths 3824 and 3826 (*Reader Service No. 108*). The A.B.C.+ was developed with the multi-purpose needs of biologists in mind: colorimetry, enzymology and immunoassay. Fitted with built-in software which can perform colorimetric and enzymological dosages,

ADVERTISEMENTS

Up To 100 Human Tissues On One Slide!

Get from 20 to 100 normal or tumor tissues on **ONE** slide for monoclonal antibody screening.

Combinations of tissues, 5 kidney, 5 breast, 5 lymph nodes etc. etc., are grouped on each slide so that a whole spectrum of tumor or normal tissues are tested at one time.

These paraffin fixed tissues speed monoclonal testing, and save immeasurable time and money.

For information about **Multi-Tissue Technology** and the slides that are immediately available, call Richard Guertin at InoGene, Inc.

InoGene, Inc.

4330 Barranca Pkwy./Suite 101-364
Irvine, California 92714
(714) 651-1245

Reader Service No.3



**micro injection?
micro dissection?
micro surgery?
micro positioning?**



**SINGER
MK I**

Singer Instrument Co. Ltd.
Treborough Lodge
Roadwater, Watchet
Somerset, England, TA23 0QL
Tel: Washford (0984) 40226
Telex: 337492 Comcab G

Reader Service No.50

the unit also performs immunoenzymatic tests in the fields of kinetics and endpoint analysis, and prints and stores calibration curves. Technical features include a choice of up to eight standards, LIN, LOG, LOGIT, and 1/absorbance X and Y axes, first, second and third degree regression calculations and spline function, continuous curve printing in enzyme kinetics and keyboard programming.

Perkin-Elmer will be introducing a new low-cost **automated atomic absorption**



Perkin-Elmer's atomic absorption spectrometer.

spectrometer in booth 4025 that is capable of performing flame, graphite furnace and mercury/hydride analyses (*Reader Service No. 109*). The \$16,000–20,000 (US) Model 1100B contains a high energy, stable optical system that features computer control of all wavelength and slit settings. The 1100B and all major accessories, including the new HGA-700 graphite furnace and AS-70 furnace auto-sampler, are entirely controlled by the instrument computer, giving the user a system totally automated for graphite furnace atomic absorption. The 1100B's interface is user friendly and contains soft keys that aid the analyst in the selection of various options.

Hitachi has a scanning **double-beam UV-visible spectrophotometer**, the model U-2000 (*Reader Service No. 110*). Priced at less than \$7,000 (US), the model U-



Double-beam spectrophotometry from Hitachi.

2000 is an analog output instrument with a 9-inch CRT and non-volatile storage of methods and spectral data. Hitachi's spectrophotometer can perform wavelength programming, quantitative analysis using quadratic, linear or segment fit, smoothing, arithmetic functions, first-to fourth-derivative calculations, time scan and multicomponent analysis. The instrument has scan speeds of up to 2,400 nm min⁻¹, and prints out hard copies of results on a built-in printer/plotter. Accessories available include automated samplers, thermostatted cell holders, and

specular reflectance and polarizer devices. Hitachi will be exhibiting in booths 1473–1479.

Analyte Corporation has an **atomic absorption spectrophotometer**, the Analyte 16F, that it says is capable of determining up to 20 elements in a single liquid sample in under one minute (*Reader Service No. 111*). The \$70,000 (US) Analyte 16F uses a conventional air/acetylene or nitrous oxide/acetylene burner that provides a multielement capability previously attainable only with inductively-coupled plasma spectrometers, while preserving the freedom from spectral interference and wide sample capability of flame atomic absorption. For working with solid samples, the Analyte 16F can be fitted with the Atomsources atomizer, a procedure that requires ten minutes or less. The Analyte 16F is controlled entirely from an AT-type personal computer with colour graphics. Analyte will be displaying the Analyte 16F in booths 455–457.

Accessories

Trivector, Inc. will be launching its new **Laboratory Automation System** in booth 1256 (*Reader Service No. 112*). The \$2,000–4,000 (US) Trivector system is based around an IEEE 802.3 bus, with



Trivector's laboratory automation system.

modules for collecting and analysing analog or serial data; controlling instruments or experiments; gathering information from automatic or manual sources; and storing and archiving data, results, working methods, control sequences or programs. The system provides LIMS functions with a structured database and has "transport gateways" to other computer systems, networks or LIMS. All programs, methods and control sequences can be down-loaded through the bus, and any number of PC/AT/PS terminals can be connected to the bus.

Spex Industries, in booths 1939–2038, has introduced a **dual-grating turret** for its existing line of 0.25 m Minimate spectrometers and its new 0.34 m spectrometers (*Reader Service No. 113*). The gratings can be switched with a twist of a knob, doubling the range of the instrument and providing dispersion compatible with both photomultiplier tubes and multichannel detectors. The turret can be retrofitted to

0.25 m and 0.34 m instruments now in use. Spex has priced the dual-grating turrets at \$350 (US).

The PopTop transplantable detector capsule for **gamma-ray spectroscopy** from EG & G Instruments allows the switching of a germanium detector from one configuration to another at the user's con-



The PopTop detector capsule for gamma-ray spectroscopy from EG & G.

venience (*Reader Service No. 114*). The ability to switch detectors quickly provides flexibility and saves money by eliminating the need for multiple detectors. The \$10,000–20,000 (US) PopTop detector capsule can be used on vertical, horizontal, side-looking, down-looking, "J" configuration and portable cryostats, with no pumping required. The PopTop assembly consists of a detector element in a sealed capsule, a mating cryostat, and a means of cooling to liquid nitrogen temperature. For applications where liquid nitrogen cannot be used, the PopTop can be ordered with an Electricool option for cooling with electricity. EG & G Instruments is exhibiting in booths 2743–2745.

Varian will be introducing its new Atomsources **direct solids atomizer** for atomic absorption spectroscopy in booth 3778 (*Reader Service No. 115*). The \$18,000 (US) Atomsources accessory can be used to determine all metallic elements measurable by atomic absorption, without the spectral and matrix interferences that are commonly encountered with X-ray fluorescence and arc/spark emission techniques, says Varian. The Atomsources can be used with Varian's SpectrAA and AA-75 series of spectrometers, as well as those from other manufacturers. According to Varian, the sensitivity of measurement can be controlled over several orders of magnitude down to parts per million levels, and a single measurement cycle takes less than one minute — a fraction of the time required to develop sample solutions for typical flame atomic absorption analysis. □

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.